



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 115605**

**To: James Schultz**  
**Location: rem/2d18/2c18**  
**Art Unit: 1635**  
**Monday, March 08, 2004**

**Case Serial Number: 09/646651**

**From: Beverly Shears**  
**Location: Remsen Bldg.**  
**RM 1A54**  
**Phone: 571-272-2528**

**beverly.shears@uspto.gov**

### **Search Notes**

# SEARCH REQUEST FORM

Requestor's Name: \_\_\_\_\_ Serial Number: \_\_\_\_\_  
Date: \_\_\_\_\_ Phone: \_\_\_\_\_ Art Unit: \_\_\_\_\_

## Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

## STAFF USE ONLY

Date completed: 03-08-04  
Searcher: Beverly C 2528  
Terminal time: 20  
Elapsed time: \_\_\_\_\_  
CPU time: \_\_\_\_\_  
Total time: 25  
Number of Searches: \_\_\_\_\_  
Number of Databases: 1

### Search Site

\_\_\_\_\_ STIC  
\_\_\_\_\_ CM-1  
\_\_\_\_\_ Pre-S

### Type of Search

\_\_\_\_\_ N.A. Sequence  
\_\_\_\_\_ A.A. Sequence  
\_\_\_\_\_ Structure  
\_\_\_\_\_ Bibliographic

### Vendors

\_\_\_\_\_ IG  
\_\_\_\_\_ STN  
\_\_\_\_\_ Dialog  
\_\_\_\_\_ APS  
\_\_\_\_\_ Geninfo  
\_\_\_\_\_ SDC  
\_\_\_\_\_ DARC/Questel  
☒ Other CGN

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

M protein - protein search, using sw model

run on: March 5, 2004, 10:31:00 ; Search time 182 Seconds  
(without alignments)  
488.027 Million cell updates/sec

US-09-646-651C-1

effect score: 467

sequence: 1 TKLEHLEGLINIFHQYSVR.....EFVLVTDLTAHDNIHKE 91

coring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

searched: 6019581 seqs, 976053577 residues

total number of hits satisfying chosen parameters: 6019581

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents\_AA\_Main.\*

- 1: /cgn2\_6/ptodata/2/paa/PCTUS\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep.\*
- 7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep.\*
- 8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep.\*
- 9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep.\*
- 10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep.\*
- 11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep.\*
- 12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep.\*
- 13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep.\*
- 14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep.\*
- 15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep.\*
- 16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep.\*
- 17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep.\*
- 18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep.\*
- 19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep.\*
- 20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep.\*
- 21: /cgn2\_6/ptodata/2/paa/US097A\_COMB.pep.\*
- 22: /cgn2\_6/ptodata/2/paa/US097B\_COMB.pep.\*
- 23: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep.\*
- 24: /cgn2\_6/ptodata/2/paa/US099A\_COMB.pep.\*
- 25: /cgn2\_6/ptodata/2/paa/US099B\_COMB.pep.\*
- 26: /cgn2\_6/ptodata/2/paa/US100\_COMB.pep.\*
- 27: /cgn2\_6/ptodata/2/paa/US101\_COMB.pep.\*
- 28: /cgn2\_6/ptodata/2/paa/US102\_COMB.pep.\*
- 29: /cgn2\_6/ptodata/2/paa/US103\_COMB.pep.\*
- 30: /cgn2\_6/ptodata/2/paa/US104\_COMB.pep.\*
- 31: /cgn2\_6/ptodata/2/paa/US106\_COMB.pep.\*
- 32: /cgn2\_6/ptodata/2/paa/US107\_COMB.pep.\*
- 33: /cgn2\_6/ptodata/2/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

result	Query	Match	Length	ID	Description
No.	Score	Match	Length	ID	Description

1	467	100.0	91	1	US-08-794-000-2	Sequence 2, Appli
2	467	100.0	91	11	US-08-759-913-5	Sequence 5, Appli
3	467	100.0	91	20	US-09-646-651C-1	Sequence 1, Appli
4	467	100.0	91	22	US-09-791-537-101828	Sequence 101828,
5	455	97.4	91	11	US-08-761-289-9	Sequence 9, Appli
6	455	97.4	91	20	US-09-646-264A-1	Sequence 1, Appli
7	455	97.4	91	20	US-09-646-651A-1	Sequence 1, Appli
8	455	97.4	91	20	US-09-646-651B-1	Sequence 1, Appli
9	371	79.4	92	22	US-09-791-537-99618	Sequence 99618, A
10	366	78.4	90	15	US-09-167-705-3	Sequence 3, Appli
11	366	78.4	90	23	US-09-872-185-11	Sequence 11, Appli
12	366	78.4	90	23	US-09-872-185-12	Sequence 12, Appli
13	366	78.4	90	23	US-09-872-185B-11	Sequence 11, Appli
14	366	78.4	90	23	US-09-872-185B-12	Sequence 12, Appli
15	366	78.4	90	31	US-10-666-513-3	Sequence 3, Appli
16	332	71.1	91	22	US-09-791-537-132106	Sequence 132106,
17	332	71.1	91	26	US-10-030-937-21	Sequence 21, Appli
18	332	71.1	92	1	PCT-US95-16871-2	Sequence 2, Appli
19	332	71.1	92	11	US-08-759-913-1	Sequence 1, Appli
20	332	71.1	92	11	US-08-761-289-2	Sequence 2, Appli
21	332	71.1	92	16	US-09-227-854-2	Sequence 2, Appli
22	332	71.1	92	22	US-09-791-537-137536	Sequence 137536,
23	332	71.1	92	25	US-09-958-053-24	Sequence 24, Appli
24	332	71.1	92	26	US-10-030-937-19	Sequence 19, Appli
25	332	71.1	92	26	US-10-030-937-20	Sequence 20, Appli
26	332	71.1	92	26	US-10-030-937-23	Sequence 23, Appli
27	332	71.1	92	26	US-10-077-600-2	Sequence 2, Appli
28	332	71.1	92	32	US-10-755-889-334	Sequence 334, App
29	332	71.1	92	33	US-60-440-088-334	Sequence 334, App
30	332	71.1	92	33	US-60-452-680-15625	Sequence 12625, A
31	332	71.1	92	33	US-60-453-050-7882	Sequence 7882, Ap
32	332	71.1	92	33	US-60-453-135-7882	Sequence 7882, Ap
33	332	71.1	92	33	US-60-455-444-4409	Sequence 4409, Ap
34	332	71.1	92	33	US-60-465-241-4409	Sequence 4409, Ap
35	332	71.1	92	33	US-60-466-412-7882	Sequence 7882, Ap
36	332	71.1	92	33	US-60-469-757-334	Sequence 334, App
37	332	71.1	95	22	US-09-760-443-1495	Sequence 1495, Ap
38	332	71.1	95	22	US-09-760-457-432	Sequence 432, App
39	332	71.1	95	22	US-09-760-484-588	Sequence 588, App
40	332	71.1	95	28	US-10-212-054-1495	Sequence 1495, Ap
41	332	71.1	95	28	US-10-217-527-432	Sequence 432, App
42	306	65.5	95	1	PCT-US01-08631-57941	Sequence 57941, A
43	302	64.7	81	22	US-09-791-537-13830	Sequence 13830, A
44	289	61.9	70	22	US-09-791-537-138681	Sequence 138681,
45	271	58.0	139	22	US-03-760-484-450	Sequence 450, App

ALIGNMENTS

RESULT 1  
US-08-794-000-2  
; Sequence 2, Application PC/TDB9601337  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: Metal-Containing Ribonucleotide Polypeptides  
; NUMBER OF SEQUENCES: 4  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (BPO)  
; CURRENT APPLICATION DATA:  
; CURRENT APPLICATION DATA: APPLICATION NUMBER: US 08/794,000  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/DE96/01337  
; FILING DATE: 17-JUL-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE 195 25 992.0  
; FILING DATE: 17-JUL-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE 195 30 500.0  
; FILING DATE: 18-AUG-1995

INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 91 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-794-000-2

Query Match 100.0%; Score 467; DB 1; Length 91;  
Best Local Similarity 100.0%; Pred. No. 1.8e-44;  
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKLEDHLEGIINIFHOYSVRLGHVDTLTKRELKQLITKELPNTLKNKTKDQGTIDKIFQNL 60  
DB 1 TKLEDHLEGIINIFHOYSVRLGHVDTLTKRELKQLITKELPNTLKNKTKDQGTIDKIFQNL 60

QY 61 DANQDEQVSFKFVVLVTDVLIITAHDNHKE 91  
DB 61 DANQDEQVSFKFVVLVTDVLIITAHDNHKE 91

RESULT 2  
US-08-759-913-5  
Sequence 5, Application US/08759913  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Goli, Surya K.  
APPLICANT: Hillman, Jennifer L.  
TITLE OF INVENTION: NOVEL HUMAN S100 PROTEINS  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: US  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/759,913  
FILING DATE: Filed Herewith  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:

ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0172 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 91 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 461678  
US-08-759-913-5

Query Match 100.0%; Score 467; DB 11; Length 91;  
Best Local Similarity 100.0%; Pred. No. 1.8e-44;  
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKLEDHLEGIINIFHOYSVRLGHVDTLTKRELKQLITKELPNTLKNKTKDQGTIDKIFQNL 60

DB 1 TKLEDHLEGIINIFHOYSVRLGHVDTLTKRELKQLITKELPNTLKNKTKDQGTIDKIFQNL 60  
QY 61 DANQDEQVSFKFVVLVTDVLIITAHDNHKE 91  
DB 61 DANQDEQVSFKFVVLVTDVLIITAHDNHKE 91

RESULT 3  
US-09-646-651C-1  
Sequence 1, Application US/09646651C  
GENERAL INFORMATION:  
APPLICANT: Kieseewetter, Stefan  
APPLICANT: Kuhn, Eckehard  
APPLICANT: Koch-Pelster, Brigitte  
APPLICANT: Brunner, Herwig  
TITLE OF INVENTION: METAL-CONTAINING RIBONUCLEOTIDE POLYPEPTIDES  
FILE REFERENCE: 206579  
CURRENT APPLICATION NUMBER: US/09/646,651C  
CURRENT FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: PCT/EP98/07722  
PRIOR FILING DATE: 1998-11-30  
PRIOR APPLICATION NUMBER: DE 198 11 047.2  
PRIOR FILING DATE: 1998-03-13  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 1  
LENGTH: 91  
TYPE: PRT  
ORGANISM: Unknown  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: ( )  
OTHER INFORMATION: Angiotropin-related protein  
US-09-646-651C-1

Query Match 100.0%; Score 467; DB 20; Length 91;  
Best Local Similarity 100.0%; Pred. No. 1.8e-44;  
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKLEDHLEGIINIFHOYSVRLGHVDTLTKRELKQLITKELPNTLKNKTKDQGTIDKIFQNL 60  
DB 1 TKLEDHLEGIINIFHOYSVRLGHVDTLTKRELKQLITKELPNTLKNKTKDQGTIDKIFQNL 60

QY 61 DANQDEQVSFKFVVLVTDVLIITAHDNHKE 91  
DB 61 DANQDEQVSFKFVVLVTDVLIITAHDNHKE 91

RESULT 4  
US-09-791-537-101828  
Sequence 101828, Application US/09791537  
GENERAL INFORMATION:  
APPLICANT: Bionomix, Inc.  
APPLICANT: Debe, Derek  
APPLICANT: Danzer, Joseph  
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE  
TITLE OF INVENTION: METHODS OF USE THEREOF  
FILE REFERENCE: 261/210  
CURRENT APPLICATION NUMBER: US/09/791,537  
CURRENT FILING DATE: 2001-02-22  
NUMBER OF SEQ ID NOS: 153055  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 101828  
LENGTH: 91  
TYPE: PRT  
ORGANISM: Sus scrofa  
US-09-791-537-101828

Query Match 100.0%; Score 467; DB 22; Length 91;  
Best Local Similarity 100.0%; Pred. No. 1.8e-44;  
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



Y 1 TKLEHLEGIINIFHOYSVRLGHVDTLTKRELKQLITKELPNTLTKNTKDKQGTIDKIFQNL 60  
b 1 TKLEHLEGIINIFHOYSVRLGHVDTLTKRELKQLITKELPNTLTKNTKDKQGTIDKIFQNL 60  
Y 61 DANODEQVSFKFVVLVTDVLITAHNDNIHKE 91  
b 61 DANODEQVSFKFVVLVTDVLITAHNDNIHKE 91

## ESULT 5

S-08-761-289-9  
Sequence 9, Application US/08761289  
GENERAL INFORMATION:  
APPLICANT: NI, J., ET AL.  
TITLE OF INVENTION: Human Chemotactic Cytokine I  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
ADDRESSEE: CECCHI, STEWART & OLSTEIN  
STREET: 6 BECKER FARM ROAD  
CITY: ROSELAND  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07068

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 INCH DISKETTE  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/761,289  
FILING DATE: December 6, 1996  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/008387  
FILING DATE: December 8, 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: MULLINS, J.G.  
REGISTRATION NUMBER: 33,073  
REFERENCE/DOCKET NUMBER: 325800-506 (PF210)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 91 AMINO ACIDS  
TYPE: AMINO ACID  
STRANDEDNESS:

TOPOLOGY: LINEAR  
MOLECULE TYPE: PEPTIDE

## S-08-761-289-9

Query Match 97.4%; Score 455; DB 11; Length 91;  
Best Local Similarity 97.8%; Pred. No. 4.1e-43;  
Matches 89; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Y 1 TKLEHLEGIINIFHOYSVRLGHVDTLTKRELKQLITKELPNTLTKNTKDKQGTIDKIFQNL 60  
b 1 TKLEHLEGIINIFHOYSVRLGHVDTLTKRELKQLITKELPNTLTKNTKDKQGTIDKIFQNL 60  
Y 61 DANODEQVSFKFVVLVTDVLITAHNDNIHKE 91  
b 61 DANODEQVSFKFVVLVTDVLITAHNDNIHKE 91

## ESULT 6

S-09-646-264A-1  
Sequence 1, Application US/09646264A  
GENERAL INFORMATION:  
APPLICANT: Klesewetter, Stefan  
APPLICANT: Seibt, Jorg-Volker  
APPLICANT: Noll, Michaela  
TITLE OF INVENTION: METAL-CONTAINING RIBONUCLEOTIDE POLYPEPTIDES

FILE REFERENCE: 206604  
CURRENT APPLICATION NUMBER: US/09/646,264A  
PRIOR FILING DATE: 2001-05-07  
PRIOR APPLICATION NUMBER: PCT/EP99/00834  
PRIOR FILING DATE: 1999-02-09  
PRIOR APPLICATION NUMBER: DE19810998.9  
PRIOR FILING DATE: 1998-03-13  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 1  
LENGTH: 91  
TYPE: PRT  
ORGANISM: Unknown  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: ().()  
OTHER INFORMATION: Angiotropin-related protein  
US-09-646-264A-1

Query Match 97.4%; Score 455; DB 20; Length 91;  
Best Local Similarity 97.8%; Pred. No. 4.1e-43;  
Matches 89; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TKLEHLEGIINIFHOYSVRLGHVDTLTKRELKQLITKELPNTLTKNTKDKQGTIDKIFQNL 60  
Db 1 TKLEHLEGIINIFHOYSVRLGHVDTLTKRELKQLITKELPNTLTKNTKDKQGTIDKIFQNL 60  
QY 61 DANODEQVSFKFVVLVTDVLITAHNDNIHKE 91  
Db 61 DANODEQVSFKFVVLVTDVLITAHNDNIHKE 91

## RESULT 7

US-09-646-651A-1  
Sequence 1, Application US/09646651A  
GENERAL INFORMATION:  
APPLICANT: Klesewetter, Stefan  
APPLICANT: Kuhn, Eckehard  
APPLICANT: Koch-Pelster, Brigitte  
APPLICANT: Brunner, Herwig  
TITLE OF INVENTION: METAL-CONTAINING RIBONUCLEOTIDE POLYPEPTIDES  
FILE REFERENCE: 206579  
CURRENT APPLICATION NUMBER: US/09/646,651A  
CURRENT FILING DATE: 2000-09-13  
PRIOR APPLICATION NUMBER: PCT/EP98/07722  
PRIOR FILING DATE: 1998-11-30  
PRIOR APPLICATION NUMBER: DE 198 11 047.2  
PRIOR FILING DATE: 1998-03-13  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 1  
LENGTH: 91  
TYPE: PRT  
ORGANISM: Unknown  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: ().()  
OTHER INFORMATION: Angiotropin-related protein  
US-09-646-651A-1

Query Match 97.4%; Score 455; DB 20; Length 91;  
Best Local Similarity 97.8%; Pred. No. 4.1e-43;  
Matches 89; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TKLEHLEGIINIFHOYSVRLGHVDTLTKRELKQLITKELPNTLTKNTKDKQGTIDKIFQNL 60  
Db 1 TKLEHLEGIINIFHOYSVRLGHVDTLTKRELKQLITKELPNTLTKNTKDKQGTIDKIFQNL 60  
QY 61 DANODEQVSFKFVVLVTDVLITAHNDNIHKE 91  
Db 61 DANODEQVSFKFVVLVTDVLITAHNDNIHKE 91

RESULT 8  
US-09-646-651B-1  
; Sequence 1, Application US/09646651B  
; GENERAL INFORMATION:  
; APPLICANT: Klesewetter, Stefan  
; APPLICANT: Kuhn, Eckehard  
; APPLICANT: Koch-Pelster, Brigitte  
; APPLICANT: Brunner, Herwig  
; TITLE OF INVENTION: METAL-CONTAINING RIBONUCLEOTIDE POLYPEPTIDES  
; FILE REFERENCE: 206579  
; CURRENT APPLICATION NUMBER: US/09/646,651B  
; CURRENT FILING DATE: 2000-09-13  
; PRIOR APPLICATION NUMBER: PCT/EP98/07722  
; PRIOR FILING DATE: 1998-11-30  
; PRIOR APPLICATION NUMBER: DE 198 11 047.2  
; PRIOR FILING DATE: 1998-03-13  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 1  
; LENGTH: 91  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (...)  
; OTHER INFORMATION: Angiotropin-related protein  
US-09-646-651B-1  
Query Match 97.4%; Score 455; DB 20; Length 91;  
Best Local Similarity 97.8%; Pred. No. 4.1e-43;  
Matches 89; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 TKLEDHLEGIINIFHOYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDOGTIDKIFQNL 60  
DB 1 TKLEDHLEGIINIFHOYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDOGTIDKIFYNL 60  
QY 61 DANODEQVSFKFEFVVLVTDVLITAHDNHKE 91  
DB 61 DANODEQVSFKFEFVVLVTDVLITAHDNHKE 91  
RESULT 9  
US-09-791-537-99618  
; Sequence 99618, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 99618  
; LENGTH: 92  
; TYPE: PRT  
; ORGANISM: Bos taurus  
US-09-791-537-99618  
Query Match 79.4%; Score 371; DB 22; Length 92;  
Best Local Similarity 81.3%; Pred. No. 1.3e-33;  
Matches 74; Conservative 9; Mismatches 8; Indels 0; Gaps 0;  
QY 1 TKLEDHLEGIINIFHOYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDOGTIDKIFQNL 60  
DB 2 TKLEDHLEGIINIFHOYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDOGTIDKIFQDL 61  
QY 61 DANODEQVSFKFEFVVLVTDVLITAHDNHKE 91  
DB 62 DADKDGAVSFEEFVVLVSRVLKTAHIDIHKE 92

RESULT 10  
US-09-167-705-3  
; Sequence 3, Application US/09167705B  
; GENERAL INFORMATION:  
; APPLICANT: Schmidt, Ann Marie  
; APPLICANT: Stern, David  
; TITLE OF INVENTION: Extracellular Novel RAGE Binding Protein (EN-RAGE) and  
; TITLE OF INVENTION: Uses Thereof  
; FILE REFERENCE: 0575/55873  
; CURRENT APPLICATION NUMBER: US/09/167,705B  
; CURRENT FILING DATE: 1998-10-06  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 90  
; TYPE: PRT  
; ORGANISM: Human  
US-09-167-705-3  
Query Match 78.4%; Score 366; DB 15; Length 90;  
Best Local Similarity 81.1%; Pred. No. 4.7e-33;  
Matches 73; Conservative 9; Mismatches 8; Indels 0; Gaps 0;  
QY 1 TKLEDHLEGIINIFHOYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDOGTIDKIFQNL 60  
DB 1 TKLEDHLEGIINIFHOYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDOGTIDKIFQDL 60  
QY 61 DANODEQVSFKFEFVVLVTDVLITAHDNHKE 90  
DB 61 DADKDGAVSFEEFVVLVSRVLKTAHIDIHKE 90  
RESULT 11  
US-09-872-185-11  
; Sequence 11, Application US/09872185  
; GENERAL INFORMATION:  
; APPLICANT: Stern, David M.  
; APPLICANT: Herold, Kevan  
; APPLICANT: Yan, Shi Du  
; APPLICANT: Schmidt, Ann Marie  
; APPLICANT: Lamster, Ira  
; TITLE OF INVENTION: METHODS FOR TREATING INFLAMMATION  
; FILE REFERENCE: 0575/64080  
; CURRENT APPLICATION NUMBER: US/09/872,185  
; CURRENT FILING DATE: 2001-08-20  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 11  
; LENGTH: 90  
; TYPE: PRT  
; ORGANISM: Bovine  
US-09-872-185-11  
Query Match 78.4%; Score 366; DB 23; Length 90;  
Best Local Similarity 81.1%; Pred. No. 4.7e-33;  
Matches 73; Conservative 9; Mismatches 8; Indels 0; Gaps 0;  
QY 1 TKLEDHLEGIINIFHOYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDOGTIDKIFQNL 60  
DB 1 TKLEDHLEGIINIFHOYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDOGTIDKIFQDL 60  
QY 61 DANODEQVSFKFEFVVLVTDVLITAHDNHKE 90  
DB 61 DADKDGAVSFEEFVVLVSRVLKTAHIDIHKE 90  
RESULT 12  
US-09-872-185-12  
; Sequence 12, Application US/09872185  
; GENERAL INFORMATION:  
; APPLICANT: Stern, David M.  
; APPLICANT: Herold, Kevan

APPLICANT: Yan, Shi Du  
APPLICANT: Schmidt, Ann Marie  
APPLICANT: Lamster, Ira  
TITLE OF INVENTION: METHODS FOR TREATING INFLAMMATION  
FILE REFERENCE: 0575/64080  
CURRENT APPLICATION NUMBER: US/09/872,185  
CURRENT FILING DATE: 2001-08-20  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 12  
LENGTH: 90  
TYPE: PRT  
ORGANISM: Bovine  
US-09-872-185-12

Query Match 78.4%; Score 366; DB 23; Length 90;  
Best Local Similarity 81.1%; Pred. No. 4.7e-33;  
Matches 73; Conservative 9; Mismatches 8; Indels 0; Gaps 0;  
Y 1 TKLEDHLEGIINIFHQYSVRLGHYDTLTKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60  
b 1 TKLEDHLEGIINIFHQYSVRLGHYDTLTKRELKQLITKELPNTLKNTKDQGTIDKIFQDL 60  
Y 61 DANODEQVSFKFEFVVLVDVLITAHDNH 90  
b 61 DADKDGAVSFEEFVVLVSRVLKTAHIDIHK 90

RESULT 13  
S-09-872-185B-11  
Sequence 11, Application US/09872185B  
GENERAL INFORMATION:  
APPLICANT: Stern, David M.  
APPLICANT: Herold, Kevan  
APPLICANT: Yan, Shi Du  
APPLICANT: Schmidt, Ann Marie  
APPLICANT: Lamster, Ira  
TITLE OF INVENTION: METHODS FOR TREATING INFLAMMATION  
FILE REFERENCE: 0575/64080  
CURRENT APPLICATION NUMBER: US/09/872,185B  
CURRENT FILING DATE: 2001-06-01  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 11  
LENGTH: 90  
TYPE: PRT  
ORGANISM: Bovine  
US-09-872-185B-11

Query Match 78.4%; Score 366; DB 23; Length 90;  
Best Local Similarity 81.1%; Pred. No. 4.7e-33;  
Matches 73; Conservative 9; Mismatches 8; Indels 0; Gaps 0;  
Y 1 TKLEDHLEGIINIFHQYSVRLGHYDTLTKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60  
b 1 TKLEDHLEGIINIFHQYSVRLGHYDTLTKRELKQLITKELPNTLKNTKDQGTIDKIFQDL 60  
Y 61 DANODEQVSFKFEFVVLVDVLITAHDNH 90  
b 61 DADKDGAVSFEEFVVLVSRVLKTAHIDIHK 90

RESULT 14  
S-09-872-185B-12  
Sequence 12, Application US/09872185B  
GENERAL INFORMATION:  
APPLICANT: Stern, David M.  
APPLICANT: Herold, Kevan  
APPLICANT: Yan, Shi Du  
APPLICANT: Schmidt, Ann Marie  
APPLICANT: Lamster, Ira  
TITLE OF INVENTION: METHODS FOR TREATING INFLAMMATION  
FILE REFERENCE: 0575/64080

CURRENT APPLICATION NUMBER: US/09/872,185B  
CURRENT FILING DATE: 2001-06-01  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 12  
LENGTH: 90  
TYPE: PRT  
ORGANISM: Bovine  
US-09-872-185B-12

Query Match 78.4%; Score 366; DB 23; Length 90;  
Best Local Similarity 81.1%; Pred. No. 4.7e-33;  
Matches 73; Conservative 9; Mismatches 8; Indels 0; Gaps 0;  
QY 1 TKLEDHLEGIINIFHQYSVRLGHYDTLTKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60  
Db 1 TKLEDHLEGIINIFHQYSVRLGHYDTLTKRELKQLITKELPNTLKNTKDQGTIDKIFQDL 60  
QY 61 DANODEQVSFKFEFVVLVDVLITAHDNH 90  
Db 61 DADKDGAVSFEEFVVLVSRVLKTAHIDIHK 90

RESULT 15  
US-10-666-513-3  
Sequence 3, Application US/10666513  
GENERAL INFORMATION:  
APPLICANT: Schmidt, Ann Marie  
APPLICANT: Stern, David  
TITLE OF INVENTION: Extracellular Novel RAGE Binding Protein (EN-RAGE) and  
TITLE OF INVENTION: Uses Thereof  
FILE REFERENCE: 0575/55873  
CURRENT APPLICATION NUMBER: US/10/666,513  
CURRENT FILING DATE: 2003-09-19  
PRIOR APPLICATION NUMBER: US/09/167,705B  
PRIOR FILING DATE: 1998-10-06  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 3  
LENGTH: 90  
TYPE: PRT  
ORGANISM: Human  
US-10-666-513-3

Query Match 78.4%; Score 366; DB 31; Length 90;  
Best Local Similarity 81.1%; Pred. No. 4.7e-33;  
Matches 73; Conservative 9; Mismatches 8; Indels 0; Gaps 0;  
QY 1 TKLEDHLEGIINIFHQYSVRLGHYDTLTKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60  
Db 1 TKLEDHLEGIINIFHQYSVRLGHYDTLTKRELKQLITKELPNTLKNTKDQGTIDKIFQDL 60  
QY 61 DANODEQVSFKFEFVVLVDVLITAHDNH 90  
Db 61 DADKDGAVSFEEFVVLVSRVLKTAHIDIHK 90

Search completed: March 5, 2004, 10:39:08  
Job time : 183 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

MM protein - protein search, using sw model

run on: March 5, 2004, 10:31:35 ; Search time 10 Seconds  
(without alignments)  
91.517 Million cell updates/sec

title: US-09-646-651C-1

perfect score: 467  
sequence: 1 TKLEDHLEGINIFHOYSVR.....EFVLTVDVLTADNIHKE 91

scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

searched: 105621 seqs, 10056811 residues

total number of hits satisfying chosen parameters: 105621

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum match 100%

Listing first 45 summaries

database : Pending Patents AA.New.\*

1: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pcp.\*  
2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pcp.\*  
3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pcp.\*  
4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pcp.\*  
5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pcp.\*  
6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pcp.\*  
7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

result No.	Score	Query Match	Length	DB ID	Description
1	366	78.4	90	6	US-10-665-867-3
2	366	78.4	90	6	US-10-665-867-4
3	206	44.1	50	6	US-10-665-867-2
4	159	34.0	89	6	US-10-624-631-32
5	151	32.3	93	6	US-10-624-631-31
6	146	31.3	98	6	US-10-624-631-28
7	134.5	28.8	95	6	US-10-487-337-4
8	134	28.7	90	6	US-10-624-631-29
9	133	28.5	99	6	US-10-624-631-33
10	129.5	27.7	97	6	US-10-487-337-8
11	127	27.2	88	6	US-10-624-631-34
12	124	26.6	495	6	US-10-767-471-717
13	109	23.3	61	6	US-10-624-631-30
14	102	21.8	104	1	PCT-US04-02188-178
15	102	21.8	104	6	US-10-764-425-178
16	78	16.7	92	7	US-60-543-011-10
17	74.5	16.0	135	6	US-10-767-701-38583
18	74.5	16.0	223	6	US-10-767-701-40040
19	72	15.4	532	6	US-10-451-467A-312
20	71	15.2	226	6	US-10-767-701-39009
21	70.5	15.1	163	6	US-10-767-701-36974
22	66.5	14.2	305	6	US-10-767-701-44602
23	66	14.1	629	6	US-10-667-290-2
24	66	14.1	629	6	US-10-667-289-2
25	65.5	14.0	498	6	US-10-767-701-46009
26	65	13.9	95	6	US-10-767-701-61761

27	64.5	13.8	198	6	US-10-767-701-61992	Sequence 61992, A
28	64.5	13.8	386	6	US-10-417-884A-5555	Sequence 5555, Ap
29	64	13.7	167	6	US-10-767-701-40255	Sequence 40255, A
30	63.5	13.6	155	6	US-10-767-701-49565	Sequence 49565, A
31	63.5	13.6	244	5	US-09-830-230A-506	Sequence 506, App
32	63.5	13.6	264	5	US-09-830-230A-505	Sequence 505, App
33	62	13.3	623	6	US-10-416-330-38	Sequence 38, Appl
34	61.5	13.2	330	6	US-10-767-701-38434	Sequence 38434, A
35	61.5	13.2	395	6	US-10-451-467A-224	Sequence 224, App
36	61	13.1	561	6	US-10-417-884A-4239	Sequence 4239, Ap
37	61	13.1	873	1	PCT-US04-03291-37	Sequence 37, Appl
38	61	13.1	873	6	US-10-771-931-37	Sequence 37, Appl
39	60.5	13.0	319	6	US-10-417-884A-4910	Sequence 4910, Ap
40	60.5	13.0	319	6	US-10-417-884A-4912	Sequence 4912, Ap
41	60	12.8	780	6	US-10-417-884A-5050	Sequence 5050, Ap
42	59.5	12.7	179	6	US-10-767-701-57231	Sequence 57231, A
43	59.5	12.7	242	6	US-10-417-884A-4507	Sequence 4507, Ap
44	59.5	12.7	242	6	US-10-417-884A-4508	Sequence 4508, Ap
45	59.5	12.7	255	6	US-10-417-884A-3995	Sequence 3995, Ap

#### ALIGNMENTS

RESULT 1  
US-10-665-867-3  
; Sequence 3, Application US/10665867  
; GENERAL INFORMATION:  
; APPLICANT: Schmidt, Ann Marie  
; APPLICANT: Stern, David  
; TITLE OF INVENTION: EXTRACELLULAR NOVEL RAGE BINDING PROTEIN (EN-RAGE) AND USES THERE-  
; FILE REFERENCE: 0575/55873-B-PCT-US  
; CURRENT APPLICATION NUMBER: US/10/665,867  
; CURRENT FILING DATE: 2003-09-17  
; PRIOR APPLICATION NUMBER: US/09/826,589  
; PRIOR FILING DATE: 2001-04-05  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 90  
; TYPE: PRT  
; ORGANISM: Bovine  
US-10-665-867-3

Query Match 78.4%; Score 366; DB 6; Length 90;  
Best Local Similarity 81.1%; Pred. No. 3.6e-34;  
Matches 73; Conservative 9; Mismatches 8; Indels 0; Gaps 0;  
QY 1 TKLEDHLEGINIFHOYSVRLGHVDTLTKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60  
DB 1 TKLEDHLEGINIFHOYSVRLGHVDTLTKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60  
QY 61 DANQBEQVSFKEFVVLVTDVLTADNIHDK 90  
DB 61 DADKDGANSFEFVVLVSRVLKTAHIDIHK 90  
RESULT 2  
US-10-665-867-4  
; Sequence 4, Application US/10665867  
; GENERAL INFORMATION:  
; APPLICANT: Schmidt, Ann Marie  
; APPLICANT: Stern, David  
; TITLE OF INVENTION: EXTRACELLULAR NOVEL RAGE BINDING PROTEIN (EN-RAGE) AND USES THERE-  
; FILE REFERENCE: 0575/55873-B-PCT-US  
; CURRENT APPLICATION NUMBER: US/10/665,867  
; CURRENT FILING DATE: 2003-09-17  
; PRIOR APPLICATION NUMBER: US/09/826,589  
; PRIOR FILING DATE: 2001-04-05  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 90

```

; TYPE: PRT
; ORGANISM: Bovine
US-10-665-867-4

Query Match      78.4%; Score 366; DB 6; Length 90;
Best Local Similarity 81.1%; Pred. No. 3.6e-34;
Matches 73; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

QY 1 TKLEDHLEGIINIFHQYSVRLGHVDTLTKRELKQITKELPNTLKNTKDQGTIDKIFQNL 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 TKLEDHLEGIINIFHQYSVRLGHVDTLTKRELKQITKELPNTLKNTKDQGTIDKIFQDL 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 DANQDEQVSPEFVVLVTDVLTIAHDNIHK 90
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 DADKGAVSEEFVVLVSRVLKTAHIDIHK 90
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 3
US-10-665-867-2
; Sequence 2, Application US/10665867
; GENERAL INFORMATION:
; APPLICANT: Schmidt, Ann Marie
; APPLICANT: Stern, David
; TITLE OF INVENTION: EXTRACELLULAR NOVEL RAGE BINDING PROTEIN (EN-PAGE) AND USES THERE
; FILE REFERENCE: 0575/55873-B-PCT-US
; CURRENT APPLICATION NUMBER: US/10/665,867
; CURRENT FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: US/09/826,589
; PRIOR FILING DATE: 2001-04-05
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (47)..(47)
; OTHER INFORMATION: x=any amino acid
US-10-665-867-2

Query Match      44.1%; Score 206; DB 6; Length 50;
Best Local Similarity 82.0%; Pred. No. 1e-16;
Matches 41; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 TKLEDHLEGIINIFHQYSVRLGHVDTLTKRELKQITKELPNTLKNTKDQ 50
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 TKLEDHLEGIINIFHQYSVRLGHVDTLTKRELKQITKELPNTLKNTKDQ 50
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 4
US-10-624-631-32
; Sequence 32, Application US/10624631
; GENERAL INFORMATION:
; APPLICANT: Morphotek, Inc.
; APPLICANT: Grasso, Luigi
; APPLICANT: Kline, J. Bradford
; APPLICANT: Nicolaides, Nicholas C.
; APPLICANT: Sasse, Philip M.
; TITLE OF INVENTION: Methods for Generating Enhanced Antibody-Producing Cell Lines wit
; FILE REFERENCE: MOR-0241/HD0002 US
; CURRENT APPLICATION NUMBER: US/10/624,631
; CURRENT FILING DATE: 2003-07-21
; PRIOR APPLICATION NUMBER: 60/397,027
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 32
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-624-631-32
```

```

Query Match      34.0%; Score 159; DB 6; Length 89;
Best Local Similarity 33.0%; Pred. No. 3.9e-11;
Matches 30; Conservative 23; Mismatches 34; Indels 4; Gaps 1;

QY 1 TKLEDHLEGIINIFHQYSVRLGHVDTLTKRELKQITKELPNTLKNTKDQGTIDKIFQNL 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3 TELEKALSNVIEVYHNYSIGIKGNHLYRDDFRKMVTECPQFVQNK---NTESLKFEL 58
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 DANQDEQVSPEFVVLVTDVLTIAHDNIHK 91
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 59 DVNSDINAINEEFLALVIRGVAAHKDSHE 89
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 5
US-10-624-631-31
; Sequence 31, Application US/10624631
; GENERAL INFORMATION:
; APPLICANT: Morphotek, Inc.
; APPLICANT: Grasso, Luigi
; APPLICANT: Kline, J. Bradford
; APPLICANT: Nicolaides, Nicholas C.
; APPLICANT: Sasse, Philip M.
; TITLE OF INVENTION: Methods for Generating Enhanced Antibody-Producing Cell Lines wit
; FILE REFERENCE: MOR-0241/HD0002 US
; CURRENT APPLICATION NUMBER: US/10/624,631
; CURRENT FILING DATE: 2003-07-21
; PRIOR APPLICATION NUMBER: 60/397,027
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 31
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-624-631-31

Query Match      32.3%; Score 151; DB 6; Length 93;
Best Local Similarity 35.2%; Pred. No. 3.2e-10;
Matches 32; Conservative 24; Mismatches 31; Indels 4; Gaps 2;

QY 1 TKLEDHLEGIINIFHQYSVRLGHVDTLTKRELKQITKELPNTLKNTKDQGTIDKIFQNL 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3 TELEKALNSIIDVHKSLIKGNPHAVYRDDLKKLLETCEPQYIRK---XGA-DWFKEL 58
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 DANQDEQVSPEFVVLVTDVLTIAHDNIHK 91
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 59 DINTDGAIVNFQEFLLIVIKMGVAAHKKSHEE 89
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 6
US-10-624-631-28
; Sequence 28, Application US/10624631
; GENERAL INFORMATION:
; APPLICANT: Morphotek, Inc.
; APPLICANT: Grasso, Luigi
; APPLICANT: Kline, J. Bradford
; APPLICANT: Nicolaides, Nicholas C.
; APPLICANT: Sasse, Philip M.
; TITLE OF INVENTION: Methods for Generating Enhanced Antibody-Producing Cell Lines wit
; FILE REFERENCE: MOR-0241/HD0002 US
; CURRENT APPLICATION NUMBER: US/10/624,631
; CURRENT FILING DATE: 2003-07-21
; PRIOR APPLICATION NUMBER: 60/397,027
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 28
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-624-631-28
```

US-10-624-631-28

Query Match 31.3%; Score 146; DB 6; Length 98;  
Best Local Similarity 35.6%; Pred. No. 1.3e-09;  
Matches 31; Conservative 18; Mismatches 38; Indels 0; Gaps 0;  
y 1 TKLEHLEGIINIFHOYSVRLGHYDTLIKRELKOLITKELPNTLKNTKDOGTIDKIFQNL 60  
b 3 TETERCIESLIAVFOYSGKDGNTQTSKTEFLSFMTLAAFTKNQKDPGVLDMMKKL 62  
y 61 DANQBOVSFKFVVLVTDVLTIAHDN 87  
b 63 DLNCDGQLDFQEFNLIGGLAIACHDS 89

RESULT 7

US-10-487-337-4

Sequence 4, Application US/10487337

GENERAL INFORMATION:

APPLICANT: Kenji OKUSE

APPLICANT: Mark BAKER

APPLICANT: Louis POON

APPLICANT: John Nicholas WOOD

APPLICANT: Misbah MALIK-HALL

TITLE OF INVENTION: SODIUM CHANNEL REGULATORS AND MODULATORS

FILE REFERENCE: 117-492 / N.86242A GCW

CURRENT APPLICATION NUMBER: US/10/487,337

CURRENT FILING DATE: 2004-02-20

PRIOR APPLICATION NUMBER: PCT/GB02/03852

PRIOR FILING DATE: 2002-08-20

PRIOR APPLICATION NUMBER: GB 0120238.1

PRIOR FILING DATE: 2001-08-20

NUMBER OF SEQ ID NOS: 55

SOFTWARE: PatentIn version 3.1

SEQ ID NO 4

LENGTH: 95

TYPE: PRT

ORGANISM: Rattus norvegicus

US-10-487-337-4

Query Match 28.8%; Score 134.5; DB 6; Length 95;  
Best Local Similarity 32.6%; Pred. No. 2.3e-08;  
Matches 28; Conservative 25; Mismatches 30; Indels 3; Gaps 1;  
y 1 TKLEHLEGIINIFHOYSVRLGHYDTLIKRELKOLITKELPNTLKNTKDOGTIDKIFQNL 60  
b 3 SOMEHMETMTLTHRFA--GEKNVLTEDLRLVLMEREPFGLNOKDPLAVDKIMKDL 59  
y 61 DANQBOVSFKFVVLVTDVLTIAHD 86  
b 60 DQCRDKGVQFSQFSLIVAGLIIACND 85

RESULT 8

US-10-624-631-29

Sequence 29, Application US/10624631

GENERAL INFORMATION:

APPLICANT: Morphotek, Inc.

APPLICANT: Grasso, Luigi

APPLICANT: Kline, J. Bradford

APPLICANT: Nicolaides, Nicholas C.

APPLICANT: Sassi, Philip M.

TITLE OF INVENTION: Methods for Generating Enhanced Antibody-Producing Cell Lines with

TITLE OF INVENTION: Improved Growth Characteristics

FILE REFERENCE: MOR-0241/HD0002 US

CURRENT APPLICATION NUMBER: US/10/624,631

CURRENT FILING DATE: 2003-07-21

PRIOR APPLICATION NUMBER: 60/397,027

PRIOR FILING DATE: 2002-07-19

NUMBER OF SEQ ID NOS: 34

SOFTWARE: PatentIn version 3.1

SEQ ID NO 29

LENGTH: 90

TYPE: PRT  
ORGANISM: Canis familiaris  
US-10-624-631-29

Query Match 28.7%; Score 134; DB 6; Length 90;  
Best Local Similarity 34.6%; Pred. No. 2.4e-08;  
Matches 27; Conservative 18; Mismatches 33; Indels 0; Gaps 0;  
y 10 IINIFHOYSVRLGHYDTLIKRELKOLITKELPNTLKNTKDOGTIDKIFQNLNANQBOVS 69  
b 4 LIAVFOYSGKDGNTQTSKTEFLSFMTLAAFTKNQKDPGVLDMMKKLINSQGLD 63  
y 70 KFEFVVLVTDVLTIAHDN 87  
b 64 FQEFNLIGGLAIACHDS 81

RESULT 9

US-10-624-631-33

Sequence 33, Application US/10624631

GENERAL INFORMATION:

APPLICANT: Morphotek, Inc.

APPLICANT: Grasso, Luigi

APPLICANT: Kline, J. Bradford

APPLICANT: Nicolaides, Nicholas C.

APPLICANT: Sassi, Philip M.

TITLE OF INVENTION: Methods for Generating Enhanced Antibody-Producing Cell Lines with

TITLE OF INVENTION: Improved Growth Characteristics

FILE REFERENCE: MOR-0241/HD0002 US

CURRENT APPLICATION NUMBER: US/10/624,631

CURRENT FILING DATE: 2003-07-21

PRIOR APPLICATION NUMBER: 60/397,027

PRIOR FILING DATE: 2002-07-19

NUMBER OF SEQ ID NOS: 34

SOFTWARE: PatentIn version 3.1

SEQ ID NO 33

LENGTH: 99

TYPE: PRT

ORGANISM: Sus scrofa

US-10-624-631-33

Query Match 28.5%; Score 133; DB 6; Length 99;  
Best Local Similarity 32.2%; Pred. No. 3.6e-08;  
Matches 29; Conservative 21; Mismatches 40; Indels 0; Gaps 0;  
y 1 TKLEHLEGIINIFHOYSVRLGHYDTLIKRELKOLITKELPNTLKNTKDOGTIDKIFQNL 60  
b 6 TETERCIESLIAVFOYSGKDGNTQTSKTEFLSFMTLAAFTKNQKDPGVLDMMKKL 65  
y 61 DANQBOVSFKFVVLVTDVLTIAHDNIHK 90  
b 66 DLNCDGQLDFQEFNLIGGLAIACHDSFIK 95

RESULT 10

US-10-487-337-8

Sequence 8, Application US/10487337

GENERAL INFORMATION:

APPLICANT: Kenji OKUSE

APPLICANT: Mark BAKER

APPLICANT: Louis POON

APPLICANT: John Nicholas WOOD

APPLICANT: Misbah MALIK-HALL

TITLE OF INVENTION: SODIUM CHANNEL REGULATORS AND MODULATORS

FILE REFERENCE: 117-492 / N.86242A GCW

CURRENT APPLICATION NUMBER: US/10/487,337

CURRENT FILING DATE: 2004-02-20

PRIOR APPLICATION NUMBER: PCT/GB02/03852

PRIOR FILING DATE: 2002-08-20

PRIOR APPLICATION NUMBER: GB 0120238.1

PRIOR FILING DATE: 2001-08-20

NUMBER OF SEQ ID NOS: 55

SOFTWARE: PatentIn version 3.1

```

; SEQ ID NO 8
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-487-337-8

Query Match      27.7%; Score 129.5; DB 6; Length 97;
(Best Local Similarity 32.6%; Pred. No. 8.5e-08;
Matches 28; Conservative 23; Mismatches 32; Indels 3; Gaps 1;

QY. 1 TKLEPHLEGIINIFHCYSVRLGHYDLIKRELKQLITKELPNTLIKNTKQOGTIDKIFQNL 60
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3 SOMEHAMETMTFTFKFAGDKG--LTKEDLRVIMKEPFGFLENQKDPFLAVDKIMKDL 59
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 DANODEQVSKFEFVVLVTDVLITAH 86
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 60 DQCRDGKVGQSFQFSLIAGLTITACND 85
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 11
US-10-624-631-34
; Sequence 34, Application US/10624631
; GENERAL INFORMATION:
; APPLICANT: Morphotek, Inc.
; APPLICANT: Grasso, Luigi
; APPLICANT: Kline, J. Bradford
; APPLICANT: Nicolaides, Nicholas C.
; APPLICANT: Sags, Philip M.
; TITLE OF INVENTION: Methods for Generating Enhanced Antibody-Producing Cell Lines with
; FILE REFERENCE: MOR-0241/HD0002 US
; CURRENT APPLICATION NUMBER: US/10/624,631
; CURRENT FILING DATE: 2003-07-21
; PRIOR APPLICATION NUMBER: 60/397,027
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 34
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Consensus Sequence
US-10-624-631-34

Query Match      27.2%; Score 127; DB 6; Length 88;
(Best Local Similarity 34.6%; Pred. No. 1.4e-07;
Matches 27; Conservative 17; Mismatches 32; Indels 2; Gaps 1;

QY 10 IINIFHCYSVRLGHYDLIKRELKQLITKELPNTLIKNTKQOGTIDKIFQNLDAODEQVS 69
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 8 LIAPFQKTAGDKG--NNLSKTEFLSFNTEAAFTRNQKOPGVLDRMKKLIDLSDDGQLD 65
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 70 KFEFVVLVTDVLITAH 87
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 66 PQEFLNLIGGLAICHDS 83
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 12
US-10-767-471-717
; Sequence 717, Application US/10767471
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001505
; CURRENT APPLICATION NUMBER: US/10/767,471
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 50231
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 717
; LENGTH: 495
; TYPE: PRT

```

```

; ORGANISM: Homo sapiens
US-10-767-471-717

Query Match      26.6%; Score 124; DB 6; Length 495;
Best Local Similarity 34.2%; Pred. No. 3.5e-06;
Matches 27; Conservative 18; Mismatches 34; Indels 0; Gaps 0;

Qy 2 KLEHLEGIINIFHOYSVRLGHYDTLTKRELKQLITKELPNTLKNKTQDGTIDKIFQNL 61
Db 3 QLLQNINGLIEFRRYARTEGNTALTRELKLEGEFADVIKPHDPATIVDEVIRLLD 62

Qy 62 ANQDEQVSFKFEFVLVTDV 80
Db 63 EDHTGTVEFKEFLVLVFKV 81

RESULT 13
US-10-624-631-30
; Sequence 30, Application US/10624631
; GENERAL INFORMATION:
; APPLICANT: Morphotek, Inc.
; APPLICANT: Grasso, Luigi
; APPLICANT: Kline, J. Bradford
; APPLICANT: Nicolaides, Nicholas C.
; APPLICANT: Sasse, Philip M.
; TITLE OF INVENTION: Methods for Generating Enhanced Antibody-Producing Cell Lines with
; TITLE OF INVENTION: Improved Growth Characteristics
; FILE REFERENCE: MOR-0241/HD0002 US
; CURRENT APPLICATION NUMBER: US/10/624,631
; CURRENT FILING DATE: 2003-07-21
; PRIOR APPLICATION NUMBER: 60/397,027
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 30
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-10-624-631-30

Query Match      23.3%; Score 109; DB 6; Length 61;
Best Local Similarity 37.3%; Pred. No. 8.4e-06;
Matches 22; Conservative 11; Mismatches 26; Indels 0; Gaps 0;

Qy 13 IFHOYSVRLGHYDTLTKRELKQLITKELPNTLKNKTQDGTIDKIFQNLDAQDEQVSFK 71
Db 3 VFQKTAGDGHSVTLTKTEFLSFMTELAFTKNQKDPGVLDRLMKKLDLNSDGLDFQ 61

RESULT 14
PCT-US04-02188-178
; Sequence 178, Application PC/TUS0402188
; GENERAL INFORMATION:
; APPLICANT: Bayer Pharmaceuticals Corporation
; APPLICANT: Eveleigh, Deepa
; APPLICANT: Bigwood, Douglas
; APPLICANT: Taylor, Ian
; TITLE OF INVENTION: EXPRESSION PROFILES FOR COLON CANCER AND METHODS OF USE
; FILE REFERENCE: 5151
; CURRENT APPLICATION NUMBER: PCT/US04/02188
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: 60/442,582
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 178
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US04-02188-178

Query Match      21.8%; Score 102; DB 1; Length 104;
Best Local Similarity 36.0%; Pred. No. 0.00011;

```

Matches 27; Conservative 18; Mismatches 26; Indels 4; Gaps 3;

3 LEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLTKNTKDGQGTIDKIFQNLDA 62

18 VERAETTLIKNFHQYSVE-GGKETLTPSELRLDVTQQLPHLMPS--NCGLEEKI-ANLGS 73

63 NQDEQVSFKFVVLV 77

74 CNDKLEFRSPFWELI 88

RESULT 15

IS-10-764-425-178

Sequence 178, Application US/10764425

GENERAL INFORMATION:

APPLICANT: Bayer Pharmaceuticals Corporation

APPLICANT: Eveleigh, Deepa

APPLICANT: Bigwood, Douglas

APPLICANT: Taylor, Ian

TITLE OF INVENTION: EXPRESSION PROFILES FOR COLON CANCER AND METHODS OF USE

FILE REFERENCE: 5151

CURRENT APPLICATION NUMBER: US/10/764,425

PRIOR FILING DATE: 2004-01-23

PRIOR FILING DATE: 2003-01-24

NUMBER OF SEQ ID NOS: 191

SOFTWARE: PatentIn version 3.2

SEQ ID NO 178

LENGTH: 104

TYPE: PRT

ORGANISM: Homo sapiens

IS-10-764-425-178

Query Match

Best Local Similarity 21.8%; Score 102; DB 6; Length 104;

Matches 27; Conservative 18; Mismatches 26; Indels 4; Gaps 3;

3 LEDHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLTKNTKDGQGTIDKIFQNLDA 62

18 VERAETTLIKNFHQYSVE-GGKETLTPSELRLDVTQQLPHLMPS--NCGLEEKI-ANLGS 73

63 NQDEQVSFKFVVLV 77

74 CNDKLEFRSPFWELI 88

Search completed: March 5, 2004, 10:39:30

Job time : 10 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

MM nucleic - nucleic search, using sw model

run on: March 8, 2004, 04:45:44 ; Search time 3135 Seconds  
(without alignments)  
1200.930 Million cell updates/sec

Title: US-09-646-651C-2

Perfect score: 107

Sequence: 1 ggaagaaunnnunnaaun.....nnnnnnnnnnuuagcag 107

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

searched: 37577330 seqs, 17593059518 residues

Total number of hits satisfying chosen parameters: 75154660

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents NA Main:\*

1: /cgn2\_6/ptodata/2/pna/US06 COMB.seq:\*

2: /cgn2\_6/ptodata/2/pna/US06 COMB.seq:\*

3: /cgn2\_6/ptodata/2/pna/US07 COMB.seq:\*

4: /cgn2\_6/ptodata/2/pna/US08 COMB.seq:\*

5: /cgn2\_6/ptodata/2/pna/US08 COMB.seq:\*

6: /cgn2\_6/ptodata/2/pna/US08 COMB.seq:\*

7: /cgn2\_6/ptodata/2/pna/US08 COMB.seq:\*

8: /cgn2\_6/ptodata/2/pna/US08 COMB.seq:\*

9: /cgn2\_6/ptodata/2/pna/US08 COMB.seq:\*

10: /cgn2\_6/ptodata/2/pna/US08 COMB.seq:\*

11: /cgn2\_6/ptodata/2/pna/US08 COMB.seq:\*

12: /cgn2\_6/ptodata/2/pna/US08 COMB.seq:\*

13: /cgn2\_6/ptodata/2/pna/US08 COMB.seq:\*

14: /cgn2\_6/ptodata/2/pna/US09 COMB.seq:\*

15: /cgn2\_6/ptodata/2/pna/US09 COMB.seq:\*

16: /cgn2\_6/ptodata/2/pna/US09 COMB.seq:\*

17: /cgn2\_6/ptodata/2/pna/US09 COMB.seq:\*

18: /cgn2\_6/ptodata/2/pna/US09 COMB.seq:\*

19: /cgn2\_6/ptodata/2/pna/US09 COMB.seq:\*

20: /cgn2\_6/ptodata/2/pna/US09 COMB.seq:\*

21: /cgn2\_6/ptodata/2/pna/US09 COMB.seq:\*

22: /cgn2\_6/ptodata/2/pna/US09 COMB.seq:\*

23: /cgn2\_6/ptodata/2/pna/US09 COMB.seq:\*

24: /cgn2\_6/ptodata/2/pna/US09 COMB.seq:\*

25: /cgn2\_6/ptodata/2/pna/US09 COMB.seq:\*

26: /cgn2\_6/ptodata/2/pna/US09 COMB.seq:\*

27: /cgn2\_6/ptodata/2/pna/US09 COMB.seq:\*

28: /cgn2\_6/ptodata/2/pna/US09 COMB.seq:\*

29: /cgn2\_6/ptodata/2/pna/US09 COMB.seq:\*

30: /cgn2\_6/ptodata/2/pna/US09 COMB.seq:\*

31: /cgn2\_6/ptodata/2/pna/US09 COMB.seq:\*

32: /cgn2\_6/ptodata/2/pna/US09 COMB.seq:\*

33: /cgn2\_6/ptodata/2/pna/US09 COMB.seq:\*

34: /cgn2\_6/ptodata/2/pna/US09 COMB.seq:\*

35: /cgn2\_6/ptodata/2/pna/US09 COMB.seq:\*

36: /cgn2\_6/ptodata/2/pna/US09 COMB.seq:\*

37: /cgn2\_6/ptodata/2/pna/US09 COMB.seq:\*

38: /cgn2\_6/ptodata/2/pna/US09 COMB.seq:\*

39: /cgn2\_6/ptodata/2/pna/US09 COMB.seq:\*

40: /cgn2\_6/ptodata/2/pna/US09 COMB.seq:\*

41: /cgn2\_6/ptodata/2/pna/US09 COMB.seq:\*

42: /cgn2\_6/ptodata/2/pna/US09 COMB.seq:\*

43: /cgn2\_6/ptodata/2/pna/US09 COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

44: /cgn2\_6/ptodata/2/pna/US100B COMB.seq:\*

45: /cgn2\_6/ptodata/2/pna/US101A COMB.seq:\*

46: /cgn2\_6/ptodata/2/pna/US101B COMB.seq:\*

47: /cgn2\_6/ptodata/2/pna/US102A COMB.seq:\*

48: /cgn2\_6/ptodata/2/pna/US102B COMB.seq:\*

49: /cgn2\_6/ptodata/2/pna/US103A COMB.seq:\*

50: /cgn2\_6/ptodata/2/pna/US103B COMB.seq:\*

51: /cgn2\_6/ptodata/2/pna/US104A COMB.seq:\*

52: /cgn2\_6/ptodata/2/pna/US104B COMB.seq:\*

53: /cgn2\_6/ptodata/2/pna/US106 COMB.seq:\*

54: /cgn2\_6/ptodata/2/pna/US107A COMB.seq:\*

55: /cgn2\_6/ptodata/2/pna/US107B COMB.seq:\*

56: /cgn2\_6/ptodata/2/pna/US6000 COMB.seq:\*

57: /cgn2\_6/ptodata/2/pna/US6001 COMB.seq:\*

58: /cgn2\_6/ptodata/2/pna/US6002 COMB.seq:\*

59: /cgn2\_6/ptodata/2/pna/US6003 COMB.seq:\*

60: /cgn2\_6/ptodata/2/pna/US6004 COMB.seq:\*

61: /cgn2\_6/ptodata/2/pna/US6005 COMB.seq:\*

62: /cgn2\_6/ptodata/2/pna/US6006 COMB.seq:\*

63: /cgn2\_6/ptodata/2/pna/US6007 COMB.seq:\*

64: /cgn2\_6/ptodata/2/pna/US6008 COMB.seq:\*

65: /cgn2\_6/ptodata/2/pna/US6009 COMB.seq:\*

66: /cgn2\_6/ptodata/2/pna/US6010 COMB.seq:\*

67: /cgn2\_6/ptodata/2/pna/US6011 COMB.seq:\*

68: /cgn2\_6/ptodata/2/pna/US6012 COMB.seq:\*

69: /cgn2\_6/ptodata/2/pna/US6013 COMB.seq:\*

70: /cgn2\_6/ptodata/2/pna/US6014 COMB.seq:\*

71: /cgn2\_6/ptodata/2/pna/US6015 COMB.seq:\*

72: /cgn2\_6/ptodata/2/pna/US6016 COMB.seq:\*

73: /cgn2\_6/ptodata/2/pna/US6017 COMB.seq:\*

74: /cgn2\_6/ptodata/2/pna/US6018 COMB.seq:\*

75: /cgn2\_6/ptodata/2/pna/US6019 COMB.seq:\*

76: /cgn2\_6/ptodata/2/pna/US6020 COMB.seq:\*

77: /cgn2\_6/ptodata/2/pna/US6021 COMB.seq:\*

78: /cgn2\_6/ptodata/2/pna/US6022 COMB.seq:\*

79: /cgn2\_6/ptodata/2/pna/US6023A COMB.seq:\*

80: /cgn2\_6/ptodata/2/pna/US6023B COMB.seq:\*

81: /cgn2\_6/ptodata/2/pna/US6024 COMB.seq:\*

82: /cgn2\_6/ptodata/2/pna/US6025 COMB.seq:\*

83: /cgn2\_6/ptodata/2/pna/US6026 COMB.seq:\*

84: /cgn2\_6/ptodata/2/pna/US6027 COMB.seq:\*

85: /cgn2\_6/ptodata/2/pna/US6028 COMB.seq:\*

86: /cgn2\_6/ptodata/2/pna/US6029 COMB.seq:\*

87: /cgn2\_6/ptodata/2/pna/US6030 COMB.seq:\*

88: /cgn2\_6/ptodata/2/pna/US6031 COMB.seq:\*

89: /cgn2\_6/ptodata/2/pna/US6032 COMB.seq:\*

90: /cgn2\_6/ptodata/2/pna/US6033 COMB.seq:\*

91: /cgn2\_6/ptodata/2/pna/US6034 COMB.seq:\*

92: /cgn2\_6/ptodata/2/pna/US6035 COMB.seq:\*

93: /cgn2\_6/ptodata/2/pna/US6036 COMB.seq:\*

94: /cgn2\_6/ptodata/2/pna/US6037 COMB.seq:\*

95: /cgn2\_6/ptodata/2/pna/US6038 COMB.seq:\*

96: /cgn2\_6/ptodata/2/pna/US6039 COMB.seq:\*

97: /cgn2\_6/ptodata/2/pna/US6040 COMB.seq:\*

98: /cgn2\_6/ptodata/2/pna/US6041 COMB.seq:\*

99: /cgn2\_6/ptodata/2/pna/US6042 COMB.seq:\*

100: /cgn2\_6/ptodata/2/pna/US6043 COMB.seq:\*

101: /cgn2\_6/ptodata/2/pna/US6044 COMB.seq:\*

102: /cgn2\_6/ptodata/2/pna/US6045 COMB.seq:\*

103: /cgn2\_6/ptodata/2/pna/US6046 COMB.seq:\*

104: /cgn2\_6/ptodata/2/pna/US6047 COMB.seq:\*

105: /cgn2\_6/ptodata/2/pna/US6048 COMB.seq:\*

106: /cgn2\_6/ptodata/2/pna/US6049 COMB.seq:\*

107: /cgn2\_6/ptodata/2/pna/US6050 COMB.seq:\*

108: /cgn2\_6/ptodata/2/pna/US6051 COMB.seq:\*

109: /cgn2\_6/ptodata/2/pna/US6052 COMB.seq:\*

110: /cgn2\_6/ptodata/2/pna/US6053 COMB.seq:\*

```
*
Result No. Score Match Length DB ID Description
-----
1 49 45.8 107 27 US-09-646-651B-2 Sequence 2, Appli
2 49 45.8 107 27 US-09-646-651C-2 Sequence 2, Appli
3 31.4 29.3 298 27 US-09-646-651A-4 Sequence 4, Appli
4 31.4 29.3 298 27 US-09-646-651B-4 Sequence 4, Appli
5 31.4 29.3 298 27 US-09-646-651C-4 Sequence 4, Appli
6 29.8 27.9 226307 103 US-60-466-412-84605 Sequence 84605, A
7 29.4 27.5 1664976 1 PCT-US97-14900A-1 Sequence 1, Appli
8 29.4 27.5 1664976 29 US-09-692-570-1 Sequence 96, Appli
9 29.2 27.3 609 69 US-60-138-831-96 Sequence 38, App
10 29.2 27.3 903 70 US-60-140-806-338 Sequence 255, App
11 29.2 27.3 2334 70 US-60-141-856-255 Sequence 20990, A
12 29.2 27.3 3020 75 US-60-173-464-20090 Sequence 25467, A
13 29.2 27.3 3020 75 US-60-191-637-25467 Sequence 20076, A
14 29.2 27.3 3020 75 US-60-191-681-20076 Sequence 270, App
15 29.2 27.3 3238 70 US-60-142-844-270 Sequence 25348, A
16 29.2 27.3 3456 25 US-09-614-150-25348 Sequence 25348, A
17 29.2 27.3 3456 25 US-09-614-150A-25348 Sequence 25336, A
18 29.2 27.3 4829 75 US-09-614-150-25336 Sequence 25336, A
19 29.2 27.3 4829 75 US-09-614-150A-25336 Sequence 20978, A
20 29.2 27.3 4829 75 US-60-173-464-20978 Sequence 25455, A
21 29.2 27.3 4829 75 US-60-191-637-25455 Sequence 20064, A
22 29.2 27.3 4829 75 US-60-191-681-20064 Sequence 1870, Ap
23 29.2 27.3 6062 72 US-60-167-217-9686 Sequence 682, App
24 29.2 27.3 61423 21 US-09-528-237A-1870 Sequence 51556, A
25 29 27.1 27614 39 US-09-948-941-682 Sequence 160, App
26 29 27.1 35807 103 US-60-465-241-51556 Sequence 3425, Ap
27 28 26.2 183928 77 US-60-212-864-160 Sequence 3425, Ap
28 27.8 26.0 510 25 US-09-605-702-3425 Sequence 280, App
29 27.8 26.0 510 53 US-10-680-765-3425 Sequence 493, App
30 27.4 25.6 474 104 US-60-475-872-2653 Sequence 493, App
31 27.4 25.6 502 104 US-60-475-872-2591 Sequence 280, App
32 27.4 25.6 524 1 PCT-US02-28214-280 Sequence 280, App
33 27.4 25.6 524 43 US-10-012-697-280 Sequence 280, App
34 27.4 25.6 524 84 US-60-275-688-280 Sequence 280, App
35 27.4 25.6 761 1 PCT-US02-28214-493 Sequence 493, App
36 27.4 25.6 761 43 US-10-012-697-493 Sequence 493, App
37 27.4 25.6 769 1 PCT-US02-28214-418 Sequence 418, App
38 27.4 25.6 769 43 US-10-012-697-418 GENERAL INFORMATI
39 27.4 25.6 769 84 US-60-275-688-419 Sequence 419, App
40 27.4 25.6 789 1 PCT-US02-28214-415 Sequence 415, App
41 27.4 25.6 789 43 US-10-012-697-415 Sequence 415, App
42 27.4 25.6 789 84 US-60-275-688-416 Sequence 416, App
43 27.4 25.6 835 84 US-60-275-688-494 Sequence 494, App
44 27.4 25.6 929 1 PCT-US02-28214-1307 Sequence 1307, Ap
45 27.4 25.6 929 43 US-10-012-697-1307 Sequence 1307, Ap

ALIGNMENTS

RESULT 1
US-09-646-651B-2
; Sequence 2, Application US/0964651B
; GENERAL INFORMATION:
; APPLICANT: Kieseletter, Stefan
; APPLICANT: Kuhn, Eckehard
; APPLICANT: Koch-Pelster, Brigitte
; APPLICANT: Brunner, Herwig
; TITLE OF INVENTION: METAL-CONTAINING RIBONUCLEOTIDE POLYPEPTIDES
; FILE REFERENCE: 206579
; CURRENT APPLICATION NUMBER: US/09/646,651B
; CURRENT FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: PCT/EP98/07722
; PRIOR FILING DATE: 1998-11-30
; PRIOR APPLICATION NUMBER: DE 198 11 047.2
; PRIOR FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
```

```
LENGTH: 107
TYPE: RNA
ORGANISM: Artificial
FEATURE:
OTHER INFORMATION: Synthetic
FEATURE:
NAME/KEY: misc_feature
LOCATION: (8)..(8)
OTHER INFORMATION: N is nucleotide A, C, U, or G
FEATURE:
NAME/KEY: misc_feature
LOCATION: (9)..(9)
OTHER INFORMATION: N is nucleotide A, C, U, or G
FEATURE:
NAME/KEY: misc_feature
LOCATION: (10)..(10)
OTHER INFORMATION: N is nucleotide A, C, U, or G
FEATURE:
NAME/KEY: misc_feature
LOCATION: (11)..(11)
OTHER INFORMATION: N is nucleotide A, C, U, or G
FEATURE:
NAME/KEY: misc_feature
LOCATION: (12)..(12)
OTHER INFORMATION: N is nucleotide A, C, U, or G
FEATURE:
NAME/KEY: misc_feature
LOCATION: (14)..(14)
OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
FEATURE:
NAME/KEY: misc_feature
LOCATION: (20)..(20)
OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
FEATURE:
NAME/KEY: misc_feature
LOCATION: (21)..(21)
OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
FEATURE:
NAME/KEY: misc_feature
LOCATION: (22)..(22)
OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
FEATURE:
NAME/KEY: misc_feature
LOCATION: (23)..(23)
OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
FEATURE:
NAME/KEY: misc_feature
LOCATION: (24)..(24)
OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
FEATURE:
NAME/KEY: misc_feature
LOCATION: (25)..(25)
OTHER INFORMATION: N is nucleotide A, C, U, or G and may or may not be present
FEATURE:
NAME/KEY: misc_feature
LOCATION: (28)..(28)
OTHER INFORMATION: N is nucleotide A, C, U, or G
FEATURE:
NAME/KEY: misc_feature
LOCATION: (29)..(29)
OTHER INFORMATION: N is nucleotide A, C, U, or G
FEATURE:
NAME/KEY: misc_feature
LOCATION: (30)..(30)
OTHER INFORMATION: N is nucleotide A, C, U, or G
FEATURE:
NAME/KEY: misc_feature
LOCATION: (34)..(34)
OTHER INFORMATION: N is nucleotide A, C, U, or G
FEATURE:
NAME/KEY: misc_feature
LOCATION: (35)..(35)
OTHER INFORMATION: N is nucleotide A, C, U, or G
```







```
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 298
; TYPE: RNA
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(298)
; OTHER INFORMATION: snRNA
US-09-646-651C-4
Query Match      29.3%; Score 31.4; DB 27; Length 298;
Best Local Similarity 42.3%; Pred. No. 58;
Matches 32; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 32 UUNNNNNAAAAAUAUAAACAUNNNNNCUUAGNNNNNNNNNNNNNNNNNAGAAUUNNNNN 91
Db 89 UUGUACAUAUAAUAAUAAACAUGACUUCUUUAGACACUCCUUAUAGAAUAAUAAUAA 148

QY 92 NNNNNNNUUAGCAG 107
Db 149 AAUAACAUAUAGCAG 164

RESULT 6
US-60-466-412-84605
; Sequence 84605, Application US/60466412
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001466
; CURRENT APPLICATION NUMBER: US/60/466,412
; CURRENT FILING DATE: 2003-04-30
; NUMBER OF SEQ ID NOS: 429241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 84605
; LENGTH: 226307
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(226307)
; OTHER INFORMATION: n = A,T,C or G
US-60-466-412-84605
Query Match      27.9%; Score 29.8; DB 103; Length 226307;
Best Local Similarity 23.4%; Pred. No. 9,5e+02;
Matches 25; Conservative 12; Mismatches 70; Indels 0; Gaps 0;

QY 1 GGAAGAAUUNNNUNAUAGUNNNNNUNNNUNNNUNNNUNNNNNAAAAUAAUAAACAUNNNNN 60
Db 90594 GGAACATCATTAATTGTGACCTTTTACATGCCAGAAAATTGTGAAGCATGTCTC 90653

QY 61 UUNAGNNNNNNNNNNNAGAAUUNNNNNNNNNNNNNNNNNNNNNNNUUAGCAG 107
Db 90654 TGAATAGGCTAAAGAGAAGAAATGGAAGGAATGGCCTTTGTAG 90700

RESULT 7
PCT-US97-14900A-1/c
; Sequence 1, Application PC/TUS9714900A
; GENERAL INFORMATION:
; APPLICANT: THE INSTITUTE FOR GENOMIC RESEARCH
; APPLICANT: 9712 MEDICAL CENTER DRIVE
; APPLICANT: ROCKVILLE, MD 20850
; APPLICANT: UNITED STATES OF AMERICA
; APPLICANT: MICROBIOLOGY DEPARTMENT
; APPLICANT: CHAMPAIGN-URBANA, IL 61801
; APPLICANT: UNITED STATES OF AMERICA
; APPLICANT: SCHOOL OF MEDICINE
```

```
; APPLICANT: DEPARTMENT OF MOLECULAR BIOLOGY AND GENETICS
; APPLICANT: BALTIMORE, MD 21205
; APPLICANT: UNITED STATES OF AMERICA
; APPLICANT: APPLICANTS/INVENTORS: Bult, Carol J.
; APPLICANT: White, Owen R.
; APPLICANT: Smith, Hamilton O.
; APPLICANT: Woese, Carl R.
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic
; TITLE OF INVENTION: Archaeon, Methanococcus jannaschii
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 New York Avenue, NW, Suite 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/14900A
; FILING DATE: 22-AUG-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,428
; FILING DATE: 22-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Eric K. Steffe
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.029PC01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1664976 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
PCT-US97-14900A-1
Query Match      27.5%; Score 29.4; DB 1; Length 1664976;
Best Local Similarity 24.8%; Pred. No. 2e+03;
Matches 26; Conservative 10; Mismatches 69; Indels 0; Gaps 0;

QY 3 AAAAUNNNNNUNAUAGUNNNNNUNNNUNNNUNNNUNNNNNAAAAUAAUAAACAUNNNNNCUU 62
Db 573299 AAAATTAGTATTAGAGCTATACTATTATATGCCAAAAATATAACATTTCTGGTG 573240

QY 63 NAGNNNNNNNNNNNAGAAUUNNNNNNNNNNNNNNNNNNNNNNNUUAGCAG 107
Db 573239 AAATAATGACAGGCAAGTATTCTTAGTAGGACGAGCCGGAG 573195

RESULT 8
US-09-692-570-1/c
; Sequence 1, Application US/09692570
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275C1
; CURRENT APPLICATION NUMBER: US/09/692,570
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; PRIOR APPLICATION NUMBER: US 08/916,421
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 20
```

SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1  
LENGTH: 1664976  
TYPE: DNA  
ORGANISM: Methanococcus jannaschii  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (28222)..(28222)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (28257)..(28258)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (84773)..(84773)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (84808)..(84808)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (98120)..(98120)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (98159)..(98159)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (98239)..(98239)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (98266)..(98266)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (98343)..(98343)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (103998)..(103998)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (148948)..(148948)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (163385)..(163385)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (191989)..(191989)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (191995)..(191995)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (231980)..(231980)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (234187)..(234187)  
OTHER INFORMATION: n equals a, t, c, or g

FEATURE:  
NAME/KEY: misc feature  
LOCATION: (234220)..(234220)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (234814)..(234814)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (309398)..(309398)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (309418)..(309418)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (312837)..(312837)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (312993)..(312993)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (319226)..(319226)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (559167)..(559167)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (559241)..(559241)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (600992)..(600992)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (622708)..(622708)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (657081)..(657081)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (657203)..(657203)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (674435)..(674435)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (682442)..(682442)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (713652)..(713652)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (741684)..(741684)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (779455)..(779455)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:

Query Match	27.3%;	Score 29.2;	DB 69;	Length 609;
Best Local Similarity	25.0%;	Pred. No. 2.9e+02;		
Matches	21;	Conservative 13;	Mismatches 50;	Indels 0;
Gaps	0;			
Qy	1	GGAAAUNNNNUNUAUGNNNNNNCUNNNUUUNNNNNNAAAAAUAUAAACAUNNNNC	60	
Db	263	GGAAATAAACCCTAGTATTACAACTCTAGTTTGTACCAAAAAATAAAATCTGTATAA	204	
Qy	61	UUNAGNNNNNNNNNNNAGAAU	84	
Db	203	TTAATTTAAATTCAAATTTAAAT	180	
RESULT 10				
US-60-140-806-338				
; Sequence 338, Application US/60140806				
; GENERAL INFORMATION:				
; APPLICANT: Kerlavage, Anthony				
; TITLE OF INVENTION: ISOLATED KINASE PROTEINS, NUCLEIC ACID				
; TITLE OF INVENTION: MOLECULES ENCODING KINASE PROTEINS, AND USES THEREOF				
; FILE REFERENCE: CL000038				
; CURRENT APPLICATION NUMBER: US/60/140,806				
; CURRENT FILING DATE: 1999-06-25				
; NUMBER OF SEQ ID NOS: 1196				
; SOFTWARE: FastSeq for Windows Version 3.0				
; SEQ ID NO 338				
; LENGTH: 903				
; TYPE: DNA				
; ORGANISM: Drosophila				
US-60-140-806-338				
Query Match	27.3%;	Score 29.2;	DB 70;	Length 903;
Best Local Similarity	25.0%;	Pred. No. 3.2e+02;		
Matches	21;	Conservative 13;	Mismatches 50;	Indels 0;
Gaps	0;			
Qy	1	GGAAAUNNNNUNUAUGNNNNNNCUNNNUUUNNNNNNAAAAAUAUAAACAUNNNNC	60	
Db	641	GGAAATAAACCCTAGTATTACAACTCTAGTTTGTACCAAAAAATAAAATCTGTATAA	700	
Qy	61	UUNAGNNNNNNNNNNNAGAAU	84	
Db	701	TTAATTTAAATTCAAATTTAAAT	724	
RESULT 11				
US-60-141-856-255				
; Sequence 255, Application US/60141856				
; GENERAL INFORMATION:				
; APPLICANT: Kerlavage, Anthony				
; TITLE OF INVENTION: ISOLATED KINASE PROTEINS, NUCLEIC ACID				
; TITLE OF INVENTION: MOLECULES ENCODING KINASE PROTEINS, AND USES THEREOF				
; FILE REFERENCE: CL000044				
; CURRENT APPLICATION NUMBER: US/60/141,856				
; CURRENT FILING DATE: 1999-07-01				
; NUMBER OF SEQ ID NOS: 617				
; SOFTWARE: FastSeq for Windows Version 3.0				
; SEQ ID NO 255				
; LENGTH: 2334				
; TYPE: DNA				
; ORGANISM: Drosophila				
US-60-141-856-255				
Query Match	27.3%;	Score 29.2;	DB 70;	Length 2334;
Best Local Similarity	25.0%;	Pred. No. 4.2e+02;		
Matches	21;	Conservative 13;	Mismatches 50;	Indels 0;
Gaps	0;			
Qy	1	GGAAAUNNNNUNUAUGNNNNNNCUNNNUUUNNNNNNAAAAAUAUAAACAUNNNNC	60	
Db	2072	GGAAATAAACCCTAGTATTACAACTCTAGTTTGTACCAAAAAATAAAATCTGTATAA	2131	
Qy	61	UUNAGNNNNNNNNNNNAGAAU	84	
Db	2132	TTAATTTTAAATTCAAATTTAAAT	2155	







```
Db 232 GAAAACTCTGTTAGACATGATGTTCTATACAGAAAGAAAGATACAACTTGATGCT 291
Qy 62 UNAGNNNNNNNNNNNAGAAUUNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNUAG 104
Db 292 TCAGTAAACACATTTTAGAGATGTTGTTCTCTTTTGTGTGAG 334
```

## RESULT 2

```
US-10-779-543-22493
; Sequence 22493, Application US/10779543
; GENERAL INFORMATION:
; APPLICANT: Williams et al
; TITLE OF INVENTION: GENE PRODUCTS DIFFERENTIALLY EXPRESSED
; FILE REFERENCE: 2300-21302
; CURRENT APPLICATION NUMBER: US/10/779,543
; PRIOR FILING DATE: 2004-02-12
; PRIOR APPLICATION NUMBER: 10/076,555
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 09/217,471
; PRIOR FILING DATE: 1998-12-21
; PRIOR APPLICATION NUMBER: 60/068,755
; PRIOR FILING DATE: 1997-12-23
; PRIOR APPLICATION NUMBER: 60/080,664
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: 60/105,234
; PRIOR FILING DATE: 1998-10-21
; PRIOR APPLICATION NUMBER: 09/297,648
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: PCT/US99/01619
; TITLE OF INVENTION: IN CANCEROUS CELLS AND THEIR METHODS OF USE II
; FILE REFERENCE: 2300-21302
; CURRENT APPLICATION NUMBER: US/10/779,543
; PRIOR FILING DATE: 2004-02-12
; PRIOR APPLICATION NUMBER: 10/076,555
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 09/217,471
; PRIOR FILING DATE: 1998-12-21
; PRIOR APPLICATION NUMBER: 60/068,755
; PRIOR FILING DATE: 1997-12-23
; PRIOR APPLICATION NUMBER: 60/080,664
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: 60/105,234
; PRIOR FILING DATE: 1998-10-21
; PRIOR APPLICATION NUMBER: 09/297,648
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: PCT/US99/01619
; NAME/KEY: misc feature
; LOCATION: 534, 596, 615, 633, 641, 648, 656, 710, 729, 736, 744, 748,
; LOCATION: 749, 751, 752
; OTHER INFORMATION: n = A,T,C or G
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22493
; LENGTH: 761
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 534, 596, 615, 633, 641, 648, 656, 710, 729, 736, 744, 748,
; LOCATION: 749, 751, 752
; OTHER INFORMATION: n = A,T,C or G
; US-10-779-543-22493
```

```
Query Match 25.6%; Score 27.4; DB 6; Length 761;
Best Local Similarity 24.3%; Pred. No. 5.4;
Matches 25; Conservative 9; Mismatches 69; Indels 0; Gaps 0;

Qy 2 GAAAAUUNNNUNUAUGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCU 61
Db 237 GAAAACTCTGTTAGACATGATGTTCTATACAGAAAGAAAGATACAACTTGATGCT 296
```

```
Qy 62 UNAGNNNNNNNNNNNAGAAUUNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNUAG 104
Db 297 TCAGTAAACACATTTTAGAGATGTTGTTCTCTTTTGTGTGAG 339
```

## RESULT 3

```
US-10-779-543-22418
; Sequence 22418, Application US/10779543
; GENERAL INFORMATION:
; APPLICANT: Williams et al
; TITLE OF INVENTION: GENE PRODUCTS DIFFERENTIALLY EXPRESSED
; FILE REFERENCE: 2300-21302
```

```
; CURRENT APPLICATION NUMBER: US/10/779,543
; CURRENT FILING DATE: 2004-02-12
; PRIOR APPLICATION NUMBER: 10/076,555
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 09/217,471
; PRIOR FILING DATE: 1998-12-21
; PRIOR APPLICATION NUMBER: 60/068,755
; PRIOR FILING DATE: 1997-12-23
; PRIOR APPLICATION NUMBER: 60/080,664
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: 60/105,234
; PRIOR FILING DATE: 1998-10-21
; PRIOR APPLICATION NUMBER: 09/297,648
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: PCT/US99/01619
; PRIOR FILING DATE: 1999-01-28
; PRIOR APPLICATION NUMBER: 60/072,910
; PRIOR FILING DATE: 1998-01-28
; PRIOR APPLICATION NUMBER: 60/075,954
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/080,114
; PRIOR FILING DATE: 1998-03-31
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 23767
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22418
; LENGTH: 769
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 2, 3, 4, 8, 19, 21, 37, 40, 44, 45, 54, 57, 65, 66, 71, 500,
; LOCATION: 566, 574, 575, 576, 597, 604, 617, 633, 640, 656, 660, 680,
; LOCATION: 684, 694, 709, 716, 720, 730, 734, 750, 755, 762, 763, 766
; OTHER INFORMATION: n = A,T,C or G
; US-10-779-543-22418

Query Match 25.6%; Score 27.4; DB 6; Length 769;
Best Local Similarity 24.3%; Pred. No. 5.4;
Matches 25; Conservative 9; Mismatches 69; Indels 0; Gaps 0;

Qy 2 GAAAAUUNNNUNUAUGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCU 61
Db 88 GAAAACTCTGTTAGACATGATGTTCTATACAGAAAGAAAGATACAACTTGATGCT 147

Qy 62 UNAGNNNNNNNNNNNAGAAUUNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNUAG 104
Db 148 TCAGTAAACACATTTTAGAGATGTTGTTCTCTTTTGTGTGAG 190
```

## RESULT 4

```
US-10-779-543-22415
; Sequence 22415, Application US/10779543
; GENERAL INFORMATION:
; APPLICANT: Williams et al
; TITLE OF INVENTION: GENE PRODUCTS DIFFERENTIALLY EXPRESSED
; FILE REFERENCE: 2300-21302
; CURRENT APPLICATION NUMBER: US/10/779,543
; CURRENT FILING DATE: 2004-02-12
; PRIOR APPLICATION NUMBER: 10/076,555
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 09/217,471
; PRIOR FILING DATE: 1998-12-21
; PRIOR APPLICATION NUMBER: 60/068,755
; PRIOR FILING DATE: 1997-12-23
; PRIOR APPLICATION NUMBER: 60/080,664
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: 60/105,234
; PRIOR FILING DATE: 1998-10-21
; PRIOR APPLICATION NUMBER: 09/297,648
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: PCT/US99/01619
```

[illegible]

```
; NAME/KEY: misc feature
; LOCATION: 649, 785, 791, 795, 806, 818, 855, 858, 869, 882, 912, 916,
; LOCATION: 919, 920, 921, 923, 926, 927, 928
; OTHER INFORMATION: n = A,T,C or G
US-10-779-543-23307

Query Match      25.6%; Score 27.4; DB 6; Length 929;
Best Local Similarity 24.3%; Pred.No.5.6;
Matches 25; Conservative 9; Mismatches 69; Indels 0; Gaps 0;

QY    2   GAAAUUNNNNUAUAUGNNNNNUNNUNNUNNUNNUNNUNNUNNUNNUNNUNNUNNUNCU 61
Db    237  GAAAACTCTGTTAGACATGATGTTCCTACTAGAGAAGAAAAGATACAACATGATTGCT 296
QY    62   UNAGNNNNNNNNNNNNAGAAUUNNNNNNNNNNNNNNNNNNNNNNNUUAG 104
Db    297  TCAGTAAACACTTTTAGAGATGTGTTCCCTTTTTTGTTGAG 339

RESULT 6
US-10-775-972-126
; Sequence 126, Application US/10775972
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C21
; CURRENT APPLICATION NUMBER: US/10/775,972
; CURRENT FILING DATE: 2004-02-10
; NUMBER OF SEQ ID NOS: 563
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 126
; LENGTH: 3552
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-972-126

Query Match      23.0%; Score 24.6; DB 6; Length 3552;
Best Local Similarity 27.2%; Pred.No.43;
Matches 22; Conservative 8; Mismatches 51; Indels 0; Gaps 0;

QY    3   AAAUUNNNNUAUAUGNNNNNUNNUNNUNNUNNUNNUNNUNNUNNUNNUNNUNNUNCU 62
Db    3081  AATATCACATTATTATGTATTCACCTTTAAGTCATAGTTTAAAAATAAACAGAAATATT 3140
QY    63   NAGNNNNNNNNNNNNAGAAA 83
Db    3141  GAGTATCACTATGTGAAGAAA 3161

RESULT 7
US-10-767-471-10805/c
; Sequence 10805, Application US/10767471
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001505
; CURRENT APPLICATION NUMBER: US/10/767,471
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 50231
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10805
; LENGTH: 1790242
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1790242)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tabl
US-10-767-471-10805
```

	Query Match	23.0%;	Score 24.6;	DB 6;	Length 1790242;
	Best Local Similarity	22.9%;	Pred. No. 73;		
	Matches	24;	Conservative 9;	Mismatches 72;	Indels 0;
					Gaps 0;
Qy	3	AAAAUNNNUNNAUGNNNNNNNCUNNUUUNNNNNAAAAAUAUAACAUNNNNCCU	62		
			:	:	:
		:	:	:	:
Db	1427735	AAAATTCTCCATTGTGTTAAATGCTTTTACCAAAAAAGATGACATTGAGATT	1427676		
Qy	63	NAGNNNNNNNNNNNNVAGAAAUNNNNNNNNNNNNNNNNNNNUUAGCAG	107		
			:	:	:
Db	1427675	CAGTTCTCATTTATGTTTAAAGGGTCTATAAAAATTATGTGCAG	1427631		

[illegible]

```

RESULT 9
PCT-US04-02000-172
; Sequence 172, Application PC/TUS0402000
; GENERAL INFORMATION:
; APPLICANT: Science Applications International Corporation
; TITLE OF INVENTION: Method and System for Identifying Biological Entities in
; TITLE OF INVENTION: Biological and Environmental Samples
; FILE REFERENCE: 36609-183264 (SAIC0087-PCT)
; CURRENT APPLICATION NUMBER: PCT/US04/02000
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US 60/441,745
; PRIOR FILING DATE: 2003-01-23
; PRIOR APPLICATION NUMBER: US 60/441,806
; PRIOR FILING DATE: 2003-01-23
; NUMBER OF SEQ ID NOS: 3242
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 172
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Clostridium perfringens
PCT-US04-02000-172

Query Match 22.8%; Score 24.2; DB 1; Length 1000;
Best Local Similarity 19.4%; Pred. No. 46;

```

[illegible]

```

RESULT 10
US-10-767-471-10627/c
; Sequence 10627, Application US/10767471
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01505
; CURRENT APPLICATION NUMBER: US/10/767,471
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 50231
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10627
; LENGTH: 184809
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: misc.feature
; LOCATION: (1)...(184809)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Table
US-10-767-471-10627

```

	Query Match	22.2%;	Score 23.8;	DB 6;	Length 184809;
	Best Local Similarity	20.8%;	Pred. No. 1.2e+02;		
	Matches	21;	Conservative 10;	Mismatches 70;	Indels 0; Gaps 0;
Qy	3	AAAAUNNNNUNAUGNNNNNNNCUNNUUUUNNNNNNAAAAANUANAACAUNNNNCUU	62		
		: : :   : :       :       :       :       :			
Dd	75569	ATAAGTACAAATTATGTCGAATTTAAATTTTAATTTTAAAAAATTTAAAAAAGAAGGCTT	75530		
Qy	63	NAGNNNNNNNNNNNAGAAUNNNNNNNNNNNNNNNNUUA	103		
		:       :       :       :			
Dd	75529	TGGAGGCCCAACACATCCAAAGGAGAAAACCATCTTCAGSTA	75489		

```

RESULT 11
US-10-673-523-14/c
; Sequence 14, Application US/10673523
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: ANTISENSE MODULATION OF PHOSPHOLIPID SCRAMBLASE 4 EXPRESSION
; FILE REFERENCE: RTS-0334
; CURRENT APPLICATION NUMBER: US/10/673,523
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: US/10/012,984
; PRIOR FILING DATE: 2001-12-04
; NUMBER OF SEQ ID NOS: 92
; SEQ ID NO 14
; LENGTH: 104000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 14992
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 14993
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 14994
;

```

[illegible]

```
NAME/KEY: unsure
LOCATION: 15031
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 15032
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 15033
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 15034
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 15035
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 15036
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
LOCATION: 15037
OTHER INFORMATION: unknown
FEATURE:
NAME/KEY: unsure
```

	Query Watch	22.1%	Score 23.6;	DB 6;	Length 104800;
	Best Local Similarity	23.1%;	Pred. No. 1.3e+02;		
	Matches	24;	Conservative	8;	Mismatches 72; Indels 0; Gaps 0;
y	3	AAAAUNNNNNUNUAUGNNNNNNCUNNNUNUUUNNNNNAAAAAUNUAAAACAUNNNNNCUU	62		
			:		
b	13486	AAAAAGAAATCAAAAAGCTAAAAATTTGTTGTTTGAAGAAGATCAAAAGCTACAAACTT	13427		
			:		
y	63	NAGUNNNNNNNNNNNNNAGAAUUNNNNNNNNNNNNNNNNNNNNUUGCA	106		
			:		
	13426	TTGGGATACATCAATAAAGAAAAGAGAGAAAAGATTCAAATTACCA	13383		
			:		

RESULT 12  
 S-10-767-471-10732/c  
 Sequence 10732, Application US/10767471  
 GENERAL INFORMATION:  
 APPLICANT: CARGILL, Michele et al.  
 TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
 RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF  
 TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF  
 FILE REFERENCE: CLO01505  
 CURRENT APPLICATION NUMBER: US/10767,471  
 CURRENT FILING DATE: 2004-01-30  
 NUMBER OF SEQ ID NOS: 50231  
 SOFTWARE: FastSEQ for Windows Version 4.0  
 SEQ ID NO 10732  
 LENGTH: 138434  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 S-10-767-471-10732

Query Match	22.1%	Score 23.6;	DB 6;	Length 138434;
Best Local Similarity	23.1%	Pred. No. 1.3e+02;		
Matches	24;	Conservative	8;	Mismatches 72;
				Indels 0;
				Gaps 0;

[illegible]

```

RESULT 13
US-10-468-591A-8/c
; Sequence 8, Application US/10468591A
; GENERAL INFORMATION:
; APPLICANT: Sahin, Ugur
; APPLICANT: Tureci, Ozlem
; APPLICANT: Ludewig, Burkhard
; TITLE OF INVENTION: Method for identifying biologically active structures of microbia
; TITLE OF INVENTION: pathogens
; FILE REFERENCE: 03-757
; CURRENT APPLICATION NUMBER: US/10/468,591A
; CURRENT FILING DATE: 2004-02-06
; PRIOR APPLICATION NUMBER: PCT/EP02/01909
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: DE 10108625.1
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 1908
; TYPE: DNA
; ORGANISM: Vaccinia virus
US-10-468-591A-8

```

Query Match	21.9%	Score 23.4;	DB 6;	Length 1908;
Best Local Similarity	29.5%;	Pred. No. 85;		
Matches 18;	Conservative	9;	Mismatches 34;	Indels 0;
				Gaps 0;

Qy	2	G A A A A U N N N N T U A U G U N N N N N C U N N U U N N N N N A A A A A U A N A A A C A U N N N N C U	61
D <sub>b</sub>	1396	G A T A A G G G C A T A T A G G T C A T C A T T A T T T G G A T T A A A A T T T A G A A C A T A T A C G C G	1337
Qy	62	T 62	
D <sub>b</sub>	1336	T 1336	

```

RESULT 14
US-09-830-230A-324
; Sequence 324, Application US/09830230A
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Lyme Disease Vaccines
; FILE REFERENCE: PB481US
; CURRENT APPLICATION NUMBER: US/09/830,230A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: PCT/US98/12718
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/057,483
; PRIOR FILING DATE: 1997-09-03
; PRIOR APPLICATION NUMBER: 60/053,344
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/053,377
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/050,359
; PRIOR FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 756
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 324
; LENGTH: 3258
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-830-230A-324

```

```

Query Match      21.98; Score 23.4; DB 5; Length 3258;
Best Local Similarity 21.24; Pred. No. 92;
Matches 21; Conservative 9; Mismatches 69; Indels 0; Gaps 0;

Qy 3 AAAAUNNNNNUNUAUUGNNNNNNNCUNNNUUUNNNNNNNAAAAAUAACAAUNNNNNCUU 62
      |||||
      |||||

pb 2324 AAAAGCATTCTATAACCAAAAGGAATACGCTTATATGCTTAATATAAAAAACAAAAAGCAA 2383
      |||||
      |||||

```





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

M protein - protein search, using sw model

un on: March 5, 2004, 09:35:33 ; Search time 55 Seconds  
(without alignments)

467.488 Million cell updates/sec

title: US-09-646-651C-1

perfect score: 457

sequence: 1 TKLEDHLEGIINIFHQYSVR.....EFVLVTDVLIITAHNDNIHKE 91

coring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

searched: 1586107 seqs, 282547505 residues

total number of hits satisfying chosen parameters: 1586107

inimum DB seq length: 0

aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

atadase : A\_Geneseq\_29Jan04:\*

1: Genesep1980s:\*

2: Genesep1990s:\*

3: Genesep2000s:\*

4: Genesep2001s:\*

5: Genesep2002s:\*

6: Genesep2003as:\*

7: Genesep2003bs:\*

8: Genesep2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

result No.	Score	Query Match	Length	DB ID	Description
1	467	100.0	91	2	AAW01826 Component
2	467	100.0	91	2	AAW93819 Argiotrop
3	371	79.4	92	2	AAW03563 Calcium b
4	366	78.4	90	3	AAy90765 Bovine CA
5	366	78.4	90	3	AAy90764 Bovine co
6	332	71.1	91	4	AAb31909 Amino aci
7	332	71.1	92	2	AAW03564 Calcium b
8	332	71.1	92	2	AAW24137 Human che
9	332	71.1	92	3	AAb45542 Human S10
10	332	71.1	92	4	AAb31911 Amino aci
11	332	71.1	92	4	AAb31907 Amino aci
12	332	71.1	92	4	AAb31908 Amino aci
13	332	71.1	92	7	ADA93649 Human cal
14	306	65.5	95	4	ABG27582 Novel hum
15	213.5	45.7	114	2	AAW17062 Human mul
16	213.5	45.7	114	2	AAW60178 Human cal
17	213.5	45.7	114	2	AAy48615 Human bre
18	213.5	45.7	114	3	AAy87637 Human cal
19	213.5	45.7	114	3	AAb45539 Human S10
20	213.5	45.7	114	4	ABB44613 Human wou
21	213.5	45.7	114	4	AAb31905 Amino aci
22	213.5	45.7	114	6	ABB82712 Human MRP
23	213.5	45.7	114	7	ADB17567 Human mye
24	213.5	45.7	114	7	ADE57110 Human Pro
25	213.5	45.7	114	7	ADE34548 Human mig

26	213.5	45.7	152	4	AAW39994	Human pol
27	206	44.1	50	3	AAy90763	Human EN-
28	204.5	43.8	112	7	ADB79921	Rat intra
29	204.5	43.8	112	7	ADE57108	Rat Prote
30	204.5	43.8	113	6	ABU63335	Rat intra
31	203	43.5	115	4	AAb31930	Amino aci
32	192	41.1	46	4	ABB43183	Peptide #
33	192	41.1	46	4	AAW37021	Peptide #
34	192	41.1	46	4	ABB26281	Protein #
35	192	41.1	46	4	AAW76914	Human bon
36	192	41.1	46	4	AAW64093	Human bra
37	192	41.1	46	4	ABG58579	Human liv
38	192	41.1	46	5	ABG46027	Human pep
39	181	38.8	95	3	AAW45544	Human S10
40	181	38.8	95	3	AAW45545	Human S10
41	181	38.8	95	5	ABG77189	Prostate
42	181	38.8	95	6	ADA10980	Human cDN
43	181	38.8	113	4	ABB12007	Human Ca-
44	179	38.3	181	4	AAU31075	Novel hum
45	178.5	38.2	113	4	ABB44612	Mouse wou

#### ALIGNMENTS

##### RESULT 1

AAW01826  
ID AAW01826 standard; protein; 91 AA.

AC AAW01826;

DT 16-OCT-1997 (first entry)

DE Component of bioactive metal RNA polypeptide.

KW Bioactive; metal; RNA polypeptide; RNP; modulation; analysis;  
KW angiogenesis; vascular state; mammalian tissue; transfer; cell;  
KW genetic information; selective; alteration; nucleic acid content;  
KW leukocyte; pig; monocyto-CuRNP.

OS Sus scrofa.

XX DE19628895-A1.

XX 23-JAN-1997.

PF 17-JUL-1996; 96DE-01028895.

PR 17-JUL-1995; 95DE-01025992.

XX 18-AUG-1995; 95DE-01030500.

FA (FRAU ) FRAUNHOFER GES FOERDERUNG ANGEWANDTEN.

PI Wissler JH, Logemann E, Kieseewetter S, Heilmeyer LMG;

XX WPI; 1997-088586/09.

DR N-PSDB; AAT62569.

XX Bioactive metal RNA polypeptide - useful for modulating angiogenesis,  
etc.

PS Claim 1; Page 15; 16pp; German.

CC A novel bioactive metal RNA polypeptide (RNP) has a RNA component  
including the sequence AAT62568 and a polypeptide component having the  
sequence AAW01826, which is encoded by AAT62569. The RNP, or anti-RNP  
immunoglobulins, can be used to modulate and/or analyse angiogenesis and  
the vascular state of mammalian tissue, transfer genetic information in  
cells and selectively alter the nucleic acid content of cells. Leukocytes  
from pig's blood were cultured in medium, and the supernatant treated  
with NH4 sulphate at 35, 45 and 90% saturation to precipitate protein  
fractions. The residual supernatant was diluted to 45% NH4 sulphate  
saturation and concentrated by ultrafiltration using a 0.5 kD membrane.

CC The retenate was purified to give 8 mg of product described as monocyto-  
 CC CuRNP  
 CC Sequence 91 AA;  
 CC

Query Match 100.0%; Score 467; DB 2; Length 91;  
 Best Local Similarity 100.0%; Pred. No. 7.5e-46;  
 Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKLEHLEGIINIPHOYSVRLGHVDTLIKRELKQITKELPNTLKNTKQDGTIDKIFQNL 60  
 DB 1 TKLEHLEGIINIPHOYSVRLGHVDTLIKRELKQITKELPNTLKNTKQDGTIDKIFQNL 60

QY 61 DANQDEQVSKEFVVLVDVLTITAHNDNIHKE 91  
 DB 61 DANQDEQVSKEFVVLVDVLTITAHNDNIHKE 91

RESULT 2  
 AAW93819  
 ID AAW93819 standard; peptide; 91 AA.

XX AAW93819;  
 AC  
 DT 21-JUN-1999 (first entry)

XX Angiotropin related protein derived peptide.

XX Angiotropin related protein; ARP; ternary complex; S100 protein;  
 KW copper-containing ribonucleoprotein; copper; cell selective;  
 KW morphogenic action; blood capillary endothelial cell; confluent;  
 KW non-mitogenic induction; cell phenotype; three-dimensional organoid;  
 KW spatiotemporal supracellular organisation; chemotropic; blood vessel;  
 KW tissue neovascularisation; angiogenesis modulation.

XX Synthetic.  
 OS  
 XX DE19811047-CL.

XX 15-APR-1999.

XX 13-MAR-1998; 98DE-01011047.

XX 13-MAR-1998; 98DE-01011047.

XX (PRAU ) PRAUNHOFER GES FOERDERUNG ANGEWANDTEN.

PI Kiesewetter S, Kuhn E, Koch-Pelster B, Brunner H;

XX WPI; 1999-216114/19.

XX Copper-containing ribonucleoproteins - useful for modulating  
 XX angiogenesis.

XX Disclosure; Page 2; 16pp; German.

XX This invention describes novel copper-containing ribonucleoproteins which  
 CC are ternary complexes of an S100 protein, copper ions and RNA comprising  
 CC the following consensus sequence or its complement GGAAAUNNNUNNAUGN1-  
 CC 6CUNNUUNNNNAAN0-IUANAAACAUN0-5CUUNAGN0-13AGAA-AUN0-16UUGAGCAG where  
 CC N = G, A, U or C. The ribonucleoproteins are stated to have the following  
 CC properties (1) cell-selective morphogenic action in vitro on isolated  
 CC primary and/or cloned blood capillary endothelial cells in culture for  
 CC the non-mitogenic induction of the change in cell phenotype from the  
 CC confluent state, for non-mitogenic alteration of the spatiotemporal  
 CC supracellular organisation of cells into three-dimensional organoid,  
 CC capillary-like structures in culture, (2) a specific chemotropic action  
 CC on blood vessels in vivo, (3) induction of directional growth of blood  
 CC vessels in vivo and (4) induction of neovascularisation of tissues  
 CC through directed ingrowth of blood vessels. Their use for modulating  
 CC angiogenesis is claimed

XX Sequence 91 AA;

Query Match 100.0%; Score 467; DB 2; Length 91;  
 Best Local Similarity 100.0%; Pred. No. 7.5e-46;  
 Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKLEHLEGIINIPHOYSVRLGHVDTLIKRELKQITKELPNTLKNTKQDGTIDKIFQNL 60  
 DB 1 TKLEHLEGIINIPHOYSVRLGHVDTLIKRELKQITKELPNTLKNTKQDGTIDKIFQNL 60

QY 61 DANQDEQVSKEFVVLVDVLTITAHNDNIHKE 91  
 DB 61 DANQDEQVSKEFVVLVDVLTITAHNDNIHKE 91

RESULT 3  
 AAW03563  
 ID AAW03563 standard; protein; 92 AA.

XX AAW03563;

XX 01-MAY-1997 (first entry)

XX Calcium binding protein CAAFl.

XX Calcium binding protein; bovine; amniotic fluid; S100 protein family;  
 KW intracellular signal transduction; squamous epithelial cell; neutrophil;  
 KW macrophage; cancer; cancerous lesion; inflammation; neoplasia; cervix;  
 KW squamous cell carcinoma; skin; oesophagus; CAAFl; lung; blood disease.

XX Bos taurus.

XX EP731166-A2.

XX 11-SEP-1996.

XX 04-DEC-1995; 95EP-00119045.

XX 06-MAR-1995; 95JP-00045564.

XX 06-MAR-1995; 95JP-00070468.

XX (TOFU ) TONEN CORP.

XX (HITO/) HITOMI J.

XX Hitomi J, Yamaguchi K, Yamamura T, Kimura T;

XX WPI; 1996-403989/41.

XX N-PSDB; AAT39345.

XX New human or bovine calcium binding protein and related nucleic acid - is  
 PT a marker for inflammation, neoplasia, skin and blood diseases.

XX Claim 1; Page 21; 36pp; English.

XX This sequence represents the CAAFl calcium-binding protein isolated from  
 CC bovine amniotic fluid. CAAFl belongs to the S100 protein family, which  
 CC includes calyculin, MRP8, and MRP14. Intracellular calcium ion  
 CC concentration is one of the key factors for intracellular signal  
 CC transduction. The calcium signals are transduced by various calcium-  
 CC binding proteins, such as the protein encoded by this sequence. CAAFl is  
 CC normally expressed in squamous epithelial cells, neutrophils and  
 CC macrophages, but atypical epithelial cells are negative for CAAFl and  
 CC overexpression is observed in several types of cancer cells and  
 CC neutrophils/macrophages infiltrating cancerous lesions. Detection of  
 CC CAAFl (using antibodies in usual immunoassays) can be used to diagnose  
 CC (or monitor) inflammation, neoplasia (particularly squamous cell  
 CC carcinoma of the skin, oesophagus, lung and cervix), and skin and blood  
 CC diseases

XX Sequence 92 AA;

Query Match 79.4%; Score 371; DB 2; Length 92;  
 Best Local Similarity 81.3%; Pred. No. 8.2e-35;  
 Matches 74; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

1 TKLEHLEGIINIFHQYSVRLGHYDTLTKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60  
2 TKLEHLEGIINIFHQYSVRLGHYDTLTKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 61

61 DANQDEQVSFKFVVLVTDVLTITAHNDNIHK 91  
62 DADKDGAVSFEEFVVLVSRVLKTAHIDIHK 92

## RESULT 4

AY90765  
D D AAY90765 standard; protein; 90 AA.

XX C C AAY90765;

18-AUG-2000 (first entry)

Bovine CAAFI acid sequence SEQ ID NO:4.

Bovine; EN-RAGE; extracellular novel RAGE binding protein; receptor for advanced glycation endproduct; inflammation; inhibition; antiinflammatory; immunoglobulin; cell surface molecule; septic shock; systemic lupus erythematosus; inflammatory lupus nephritis; endotoxaemia; autoimmune disorder; inflammatory disorder.

Bos taurus.

WO200020621-A1.

13-APR-2000.

06-OCT-1999; 99WO-US023303.

06-OCT-1998; 98US-00167705.

05-MAR-1999; 99US-00263312.

(UYCO ) UNIV COLUMBIA NEW YORK.

Schmidt AM, Stern D;

WPI; 2000-303794/26.

New human EN-RAGE (extracellular novel receptor for advanced glycation end products) peptide, useful for identifying anti-inflammatory compounds that inhibit its interaction with RAGE.

Claim 2; Page 41; 132pp; English.

The present invention describes an isolated human EN-RAGE (extracellular novel receptor for advanced glycation end products) peptide (P1). The EN-RAGE peptide binds to RAGE which is a member of the immunoglobulin superfamily of cell-surface molecules. A compound capable of inhibiting the interaction of EN-RAGE with RAGE is useful for the suppression of inflammation resulting from systemic lupus erythematosus, inflammatory lupus nephritis, septic shock, endotoxaemia, or an autoimmune or inflammatory disorder in which the recruitment of EN-RAGE containing inflammatory cells occurs. The compound is also useful for the treatment of systemic lupus erythematosus, inflammatory lupus nephritis in a subject. The human EN-RAGE peptide is useful for identifying compounds that inhibit its interaction with RAGE. The present sequence represents the bovine CAAFI which shows homology to the human EN-RAGE N-terminal amino acid sequence

Sequence 90 AA;

Query Match 78.4%; Score 366; DB 3; Length 90;

Best Local Similarity 81.1%; Pred. No. 3e-34;

Matches 73; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

1 TKLEHLEGIINIFHQYSVRLGHYDTLTKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60

1 TKLEHLEGIINIFHQYSVRLGHYDTLTKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60

QY 61 DANQDEQVSFKFVVLVTDVLTITAHNDNIHK 90  
Db 61 DADKDGAVSFEEFVVLVSRVLKTAHIDIHK 90

## RESULT 5

AY90764  
ID AAY90764 standard; protein; 90 AA.

XX AC AAY90764;

18-AUG-2000 (first entry)

Bovine corneal antigen (B-COAg) acid sequence SEQ ID NO:3.

Bovine; EN-RAGE; extracellular novel RAGE binding protein; receptor for advanced glycation endproduct; inflammation; inhibition; antiinflammatory; immunoglobulin; cell surface molecule; septic shock; systemic lupus erythematosus; inflammatory lupus nephritis; endotoxaemia; autoimmune disorder; inflammatory disorder.

Bos taurus.

WO200020621-A1.

13-APR-2000.

06-OCT-1999; 99WO-US023303.

06-OCT-1998; 98US-00167705.

05-MAR-1999; 99US-00263312.

(UYCO ) UNIV COLUMBIA NEW YORK.

Schmidt AM, Stern D;

WPI; 2000-303794/26.

New human EN-RAGE (extracellular novel receptor for advanced glycation end products) peptide, useful for identifying anti-inflammatory compounds that inhibit its interaction with RAGE.

Claim 2; Page 41; 132pp; English.

The present invention describes an isolated human EN-RAGE (extracellular novel receptor for advanced glycation end products) peptide (P1). The EN-RAGE peptide binds to RAGE which is a member of the immunoglobulin superfamily of cell-surface molecules. A compound capable of inhibiting the interaction of EN-RAGE with RAGE is useful for the suppression of inflammation resulting from systemic lupus erythematosus, inflammatory lupus nephritis, septic shock, endotoxaemia, or an autoimmune or inflammatory disorder in which the recruitment of EN-RAGE containing inflammatory cells occurs. The compound is also useful for the treatment of systemic lupus erythematosus, inflammatory lupus nephritis in a subject. The human EN-RAGE peptide is useful for identifying compounds that inhibit its interaction with RAGE. The present sequence represents the bovine corneal antigen which shows homology to the human EN-RAGE N-terminal amino acid sequence

Sequence 90 AA;

Query Match 78.4%; Score 366; DB 3; Length 90;

Best Local Similarity 81.1%; Pred. No. 3e-34;

Matches 73; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

1 TKLEHLEGIINIFHQYSVRLGHYDTLTKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60

1 TKLEHLEGIINIFHQYSVRLGHYDTLTKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60

QY 61 DANQDEQVSFKFVVLVTDVLTITAHNDNIHK 90

Db 61 DADKDGAVSFEEFVVLVSRVLKTAHIDIHK 90

AC	AAW03564;
XX	
DT	01-MAY-1997 (first entry)
XX	
DE	Calcium binding protein CAAF1.
XX	
KW	Calcium binding protein; human; amniotic fluid; S100 protein family;
KW	intracellular signal transduction; squamous epithelial cell; neutrophil;
KW	macrophage; cancer; carcinosarcoma lesion; inflammation; neoplasia; cervix;
KW	squamous cell carcinoma; skin; oesophagus; CAAF1; lung; blood disease.
XX	
OS	Homo sapiens.
XX	
PN	EP731166-A2.
XX	
PD	11-SEP-1996.
XX	
PF	04-DEC-1995; 95EP-00119045.
XX	
PR	06-MAR-1995; 95JP-00045564.
XX	
PR	06-MAR-1995; 95JP-00070468.
XX	
PA	(TOFU ) TONEN CORP.
PA	(HITO/) HITOMI J.
XX	
PI	Hitomi J, Yamaguchi K, Yamamura T, Kimura T;
XX	
DR	WPI; 1996-403989/41.
DR	N-PSDB; AAT39346.
XX	
PT	New human or bovine calcium binding protein and related nucleic acid - is
PT	a marker for inflammation, neoplasia, skin and blood diseases.
XX	
PS	Claim 1; Page 24; 36pp; English.
XX	
CC	This sequence represents the CAAF1 calcium-binding protein isolated from
CC	human amniotic fluid. CAAF1 belongs to the S100 protein family, which
CC	includes calyculin, MRP8, and MRP14. Intracellular calcium ion
CC	concentration is one of the key factors for intracellular signal
CC	transduction. The calcium signals are transduced by various calcium-
CC	binding proteins, such as this protein. CAAF1 is normally expressed in
CC	squamous epithelial cells, neutrophils and macrophages, but atypical
CC	epithelial cells are negative for CAAF1 and overexpression is observed in
CC	several types of cancer cells and neutrophils/macrophages infiltrating
CC	cancerous lesions. Detection of CAAF1 (using antibodies in usual
CC	immunoassays) can be used to diagnose (or monitor) inflammation,
CC	neoplasia (particularly squamous cell carcinoma of the skin, oesophagus,
CC	lung and cervix), and skin and blood diseases
XX	
SQ	Sequence 92 AA;
Query Match	71.1%; Score 332; DB 2; Length 92;
Best Local Similarity	70.3%; Pred. No. 2.5e-30;
Matches	64; Conservative 10; Mismatches 17; Indels 0; Gaps 0;
QY	1 TKLEHLEGIIINIFHOYSVRLGHYDITLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
	:     :     :     :     :     :     :     :     :     :
Dd	2 TKLEEHLGIVNIIPHQYSVRKGHFDTLSKGELKQLTTELANTIKNIKDAVIDEIQGL 61
	:     :     :     :     :     :     :     :     :
QY	61 DANQEVSFKFVVLTVDLITAHDNHKE 91
	:     :     :     :     :     :     :     :     :
Dd	62 DANQEVDQFEFTSLVAIALKAHHYTHKE 92
	:     :     :     :     :     :     :     :     :
RESULT 8	
AAW24137	
ID	AAW24137 standard; protein; 92 AA.
XX	
AC	AAW24137;
XX	
DT	28-JAN-1998 (first entry)
AC	AAW03564;
XX	
DT	01-MAY-1997 (first entry)
XX	
DE	Calcium binding protein CAAF1.
XX	
KW	Calcium binding protein; human; amniotic fluid; S100 protein family;
KW	intracellular signal transduction; squamous epithelial cell; neutrophil;
KW	macrophage; cancer; carcinosarcoma lesion; inflammation; neoplasia; cervix;
KW	squamous cell carcinoma; skin; oesophagus; CAAF1; lung; blood disease.
XX	
OS	Homo sapiens.
XX	
PN	EP731166-A2.
XX	
PD	11-SEP-1996.
XX	
PF	04-DEC-1995; 95EP-00119045.
XX	
PR	06-MAR-1995; 95JP-00045564.
XX	
PR	06-MAR-1995; 95JP-00070468.
XX	
PA	(TOFU ) TONEN CORP.
PA	(HITO/) HITOMI J.
XX	
PI	Hitomi J, Yamaguchi K, Yamamura T, Kimura T;
XX	
DR	WPI; 1996-403989/41.
DR	N-PSDB; AAT39346.
XX	
PT	New human or bovine calcium binding protein and related nucleic acid - is
PT	a marker for inflammation, neoplasia, skin and blood diseases.
XX	
PS	Claim 1; Page 24; 36pp; English.
XX	
CC	This sequence represents the CAAF1 calcium-binding protein isolated from
CC	human amniotic fluid. CAAF1 belongs to the S100 protein family, which
CC	includes calyculin, MRP8, and MRP14. Intracellular calcium ion
CC	concentration is one of the key factors for intracellular signal
CC	transduction. The calcium signals are transduced by various calcium-
CC	binding proteins, such as this protein. CAAF1 is normally expressed in
CC	squamous epithelial cells, neutrophils and macrophages, but atypical
CC	epithelial cells are negative for CAAF1 and overexpression is observed in
CC	several types of cancer cells and neutrophils/macrophages infiltrating
CC	cancerous lesions. Detection of CAAF1 (using antibodies in usual
CC	immunoassays) can be used to diagnose (or monitor) inflammation,
CC	neoplasia (particularly squamous cell carcinoma of the skin, oesophagus,
CC	lung and cervix), and skin and blood diseases
XX	
SQ	Sequence 92 AA;
Query Match	71.1%; Score 332; DB 4; Length 91;
Best Local Similarity	70.3%; Pred. No. 2.5e-30;
Matches	64; Conservative 10; Mismatches 17; Indels 0; Gaps 0;
QY	1 TKLEHLEGIIINIFHOYSVRLGHYDITLIKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
	:     :     :     :     :     :     :     :     :
Dd	1 TKLEEHLGIVNIIPHQYSVRKGHFDTLSKGELKQLTTELANTIKNIKDAVIDEIQGL 60
	:     :     :     :     :     :     :     :     :
QY	61 DANQEVSFKFVVLTVDLITAHDNHKE 91
	:     :     :     :     :     :     :     :     :
Dd	61 DANQEVDQFEFTSLVAIALKAHHYTHKE 91
	:     :     :     :     :     :     :     :     :
RESULT 7	
AAW03564	

X Human chemotactic cytokine I.  
X chemotactic cytokine; tumour; autoimmune disease; antagonist; agonist.  
X Homo sapiens.  
X WO9723640-A1.  
X 03-JUL-1997.  
X 26-DEC-1995; 95WO-US016871.  
X 26-DEC-1995; 95WO-US016871.  
X (HUMA-) HUMAN GENOME SCI INC.  
X Ni J, Yu G, Alfonso P, Gentz R, Su JY;  
X WPI; 1997-351075/32.  
X N-PSDB; AAT85774.  
X DNA encoding chemotactic cytokine I - used to treat, e.g. tumours,  
X chronic infection, leukaemia, etc.  
X Claim 12; Page 48-49; 64pp; English.  
X This is a human chemotactic cytokine I polypeptide. The encoding  
X polynucleotide, along with a vector and a host cell can be used for the  
X recombinant production of the chemotactic cytokine. Cytokine agonists and  
X antagonists can be used for the treatment of a patient requiring a  
X chemotactic cytokine I and for the treatment of a patient requiring the  
X inhibition of a chemotactic cytokine I polypeptide, respectively. The  
X chemotactic cytokine is used to treat tumours, chronic infection,  
X leukaemia and T-cell mediated autoimmune diseases  
X  
X Sequence 92 AA;  
Query Match 71.1%; Score 332; DB 2; Length 92;  
Best Local Similarity 70.3%; Pred. No. 2.5e-30;  
Matches 64; Conservative 10; Mismatches 17; Indels 0; Gaps 0;  
Y 1 TKLDEHLEGINIFHOYSVRLGHYDTLTKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60  
b 2 TKLEHLEGINIFHOYSVRKGHFDTLKSGELKQLLTKELANIKIKDKAVIDEIPQGL 61  
Y 61 DANQEQVSFEFVVLVTDVLITAHNDNIKE 91  
b 62 DANQEQVDQFQEFISLVAIALKAAHYTHKE 92  
RESULT 9  
AB45542  
D AAB45542 standard; protein; 92 AA.  
X AAB45542;  
X 22-FEB-2001 (first entry)  
X Human S100A12 protein.  
X S100 protein; human; treatment; cardiomyopathy; cardiac insufficiency;  
X calcium-binding protein; calcium homeostasis; cardiac muscle;  
X pumping capacity; myocardial cell; systolic calcium ion release;  
X sarcoplasmic reticulum; cardiac disease; hypertension; rhythm disorder;  
X valve defect.  
X Homo sapiens.  
X DE19915485-A1.  
X 19-OCT-2000.  
X

PF 07-APR-1999; 99DE-01015485.  
XX 07-APR-1999; 99DE-01015485.  
XX (KATU/) KATUS H A.  
PA (REMP/) REMPPIS A.  
XX Katus HA, Remppis A;  
XX WPI; 2000-673510/66.  
DR N-PSDB; AAC81812.  
XX Composition containing S100 protein, corresponding nucleic acid or  
XX vector, useful for treating cardiomyopathy and cardiac insufficiency.  
XX Claim 35; Page 20; 36pp; German.  
XX This invention describes a novel composition for treating primary or  
XX secondary cardiomyopathy or cardiac insufficiency contains at least one  
XX S100 protein (I) or nucleic acid (II) encoding (I), or their mutants or  
XX fragments, or a gene transfer vector containing (II), optionally  
XX formulated with auxiliaries and/or carriers. (I) are calcium-binding  
XX proteins involved in calcium homeostasis, so their overexpression in  
XX cardiac muscle will improve pumping capacity (and overall capacity) of  
XX the heart. In cultured myocardial cells they increase the contraction and  
XX relaxation rates associated with increased systolic calcium ion release  
XX from the sarcoplasmic reticulum (SR) and calcium re-uptake by SR. (I) are  
XX used to treat cardiomyopathy (CMP) where inherited or caused by  
XX spontaneous mutations and ischemic CMP caused by arteriosclerosis,  
XX dilative CMP caused by toxic/infectious disease, cardiac disease caused  
XX by pulmonary and/or arterial hypertension, and structural disease caused  
XX by rhythm disorders or valve defects, generally any condition associated  
XX with reduced contractile force. Unlike calmodulin, which is expressed  
XX ubiquitously, (I) show tissue-specific expression and treat the  
XX underlying defect in the sarcoplasmic reticulum (SR) that causes cardiac  
XX disease  
XX  
XX Sequence 92 AA;  
Query Match 71.1%; Score 332; DB 3; Length 92;  
Best Local Similarity 70.3%; Pred. No. 2.5e-30;  
Matches 64; Conservative 10; Mismatches 17; Indels 0; Gaps 0;  
QY 1 TKLDEHLEGINIFHOYSVRLGHYDTLTKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60  
Db 2 TKLEHLEGINIFHOYSVRKGHFDTLKSGELKQLLTKELANIKIKDKAVIDEIPQGL 61  
QY 61 DANQEQVSFEFVVLVTDVLITAHNDNIKE 91  
Db 62 DANQEQVDQFQEFISLVAIALKAAHYTHKE 92  
RESULT 10  
AAB31911  
ID AAB31911 standard; protein; 92 AA.  
XX AAB31911;  
XX 15-MAY-2001 (first entry)  
XX Amino acid sequence of a human protein.  
XX Human; perlecan; retinol-binding plasma protein; calgranulin B; vaccine;  
XX ganglioside GM2 activator; saposin B; degenerative disease; glial cell;  
XX neurological disease; auto-immune disease; multiple sclerosis; toxicity;  
XX Alzheimer's disease; Parkinson's disease; myotrophic lateral sclerosis;  
XX rheumatoid polyarthritis; lupus erythematosus; gene therapy.  
XX Homo sapiens.  
XX WO2000105422-A2.  
XX 25-JAN-2001.  
PD

XX PF 17-JUL-2000; 2000WO-FR002057.  
 XX PR 15-JUL-1999; 99FR-00009372.  
 XX PR (INMR) BIOMERIEUX STELHYS.  
 XX PI Roecklin D, Kolbe H, Charles M, Malcus C, Santoro L, Perron H;  
 XX WPI; 2001-159475/16.  
 XX PT Detecting, preventing and treating degenerative, neurological and  
 XX PT autoimmune diseases, particularly multiple sclerosis, using specified  
 XX PT polypeptides or related nucleic acid or ligand.  
 XX PS Claim 1; Page 168; 209pp; French.  
 XX CC The present sequence represents a human protein, which is used in the  
 XX CC method of the invention. The specification describes a method which uses  
 XX CC at least one polypeptide or polynucleotide sequence belonging to the  
 XX CC perlecan, precursor of the retinol-binding plasma protein, precursor of  
 XX CC the ganglioside GM2 activator, calgranulin B or saposin B protein  
 XX CC families. The method is used for detecting, preventing or treating a  
 XX CC degenerative, neurological and/or auto-immune disease. The  
 XX CC polynucleotides and polypeptides are used for diagnosis, prognosis,  
 XX CC prevention and treatment of multiple sclerosis (in its various forms and  
 XX CC phases). They may also be useful in cases of e.g. Alzheimer's and  
 XX CC Parkinson's diseases, amyotrophic lateral sclerosis, rheumatoid  
 XX CC polyarthritis and lupus erythematosus, including use as vaccines and in  
 XX CC gene therapy (expression of sense or antisense sequences). They can also  
 XX CC be used to assess efficacy of potential therapeutic agents, particularly  
 XX CC compounds that reduce or inhibit toxicity towards glial cells  
 XX SQ Sequence 92 AA;  
 Query Match 71.1%; Score 332; DB 4; Length 92;  
 Best Local Similarity 70.3%; Pred. No. 2.5e-30;  
 Matches 64; Conservative 10; Mismatches 17; Indels 0; Gaps 0;  
 QY 1 TKLEHLEGIINIFHOYSVRLGHVDTLTKRELKQLTKELPNTLKNTKDQGTIDKIFQNL 60  
 Db 2 TKLEHLEGIINIFHOYSVRLGHVDTLTKRELKQLTKELPNTLKNTKDQGTIDKIFQNL 61  
 QY 61 DANDEQVSPKFEFVVLVDLITAHNDNIHKE 91  
 Db 62 DANDEQVDFQEFISLVAIALKAAHYHTHKE 92  
 RESULT 11  
 AAB31907  
 ID AAB31907 standard; protein; 92 AA.  
 XX AC AAB31907;  
 XX DT 15-MAY-2001 (first entry)  
 XX DE Amino acid sequence of a human protein.  
 XX KW Human; perlecan; retinol-binding plasma protein; calgranulin B; vaccine;  
 XX KW ganglioside GM2 activator; saposin B; degenerative disease; glial cell;  
 XX KW neurological disease; auto-immune disease; multiple sclerosis; toxicity;  
 XX KW Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;  
 XX KW rheumatoid polyarthritis; lupus erythematosus; gene therapy.  
 XX OS Homo sapiens.  
 XX PN WO200105422-A2.  
 XX PD 25-JAN-2001.  
 XX PF 17-JUL-2000; 2000WO-FR002057.  
 XX PR 15-JUL-1999; 99FR-00009372.  
 XX PR (INMR) BIOMERIEUX STELHYS.  
 XX PI Roecklin D, Kolbe H, Charles M, Malcus C, Santoro L, Perron H;

XX (INMR) BIOMERIEUX STELHYS.  
 XX PI Roecklin D, Kolbe H, Charles M, Malcus C, Santoro L, Perron H;  
 XX WPI; 2001-159475/16.  
 XX PT Detecting, preventing and treating degenerative, neurological and  
 XX PT autoimmune diseases, particularly multiple sclerosis, using specified  
 XX PT polypeptides or related nucleic acid or ligand.  
 XX PS Claim 1; Page 166-167; 209pp; French.  
 XX CC The present sequence represents a human protein, which is used in the  
 XX CC method of the invention. The specification describes a method which uses  
 XX CC at least one polypeptide or polynucleotide sequence belonging to the  
 XX CC perlecan, precursor of the retinol-binding plasma protein, precursor of  
 XX CC the ganglioside GM2 activator, calgranulin B or saposin B protein  
 XX CC families. The method is used for detecting, preventing or treating a  
 XX CC degenerative, neurological and/or auto-immune disease. The  
 XX CC polynucleotides and polypeptides are used for diagnosis, prognosis,  
 XX CC prevention and treatment of multiple sclerosis (in its various forms and  
 XX CC phases). They may also be useful in cases of e.g. Alzheimer's and  
 XX CC Parkinson's diseases, amyotrophic lateral sclerosis, rheumatoid  
 XX CC polyarthritis and lupus erythematosus, including use as vaccines and in  
 XX CC gene therapy (expression of sense or antisense sequences). They can also  
 XX CC be used to assess efficacy of potential therapeutic agents, particularly  
 XX CC compounds that reduce or inhibit toxicity towards glial cells  
 XX SQ Sequence 92 AA;  
 Query Match 71.1%; Score 332; DB 4; Length 92;  
 Best Local Similarity 70.3%; Pred. No. 2.5e-30;  
 Matches 64; Conservative 10; Mismatches 17; Indels 0; Gaps 0;  
 QY 1 TKLEHLEGIINIFHOYSVRLGHVDTLTKRELKQLTKELPNTLKNTKDQGTIDKIFQNL 60  
 Db 2 TKLEHLEGIINIFHOYSVRLGHVDTLTKRELKQLTKELPNTLKNTKDQGTIDKIFQNL 61  
 QY 61 DANDEQVSPKFEFVVLVDLITAHNDNIHKE 91  
 Db 62 DANDEQVDFQEFISLVAIALKAAHYHTHKE 92  
 RESULT 12  
 AAB31908  
 ID AAB31908 standard; protein; 92 AA.  
 XX AC AAB31908;  
 XX DT 15-MAY-2001 (first entry)  
 XX DE Amino acid sequence of a human protein.  
 XX KW Human; perlecan; retinol-binding plasma protein; calgranulin B; vaccine;  
 XX KW ganglioside GM2 activator; saposin B; degenerative disease; glial cell;  
 XX KW neurological disease; auto-immune disease; multiple sclerosis; toxicity;  
 XX KW Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;  
 XX KW rheumatoid polyarthritis; lupus erythematosus; gene therapy.  
 XX OS Homo sapiens.  
 XX PN WO200105422-A2.  
 XX PD 25-JAN-2001.  
 XX PF 17-JUL-2000; 2000WO-FR002057.  
 XX PR 15-JUL-1999; 99FR-00009372.  
 XX PR (INMR) BIOMERIEUX STELHYS.  
 XX PI Roecklin D, Kolbe H, Charles M, Malcus C, Santoro L, Perron H;

X WPI; 2001-159475/16.  
R Detecting, preventing and treating degenerative, neurological and  
X autoimmune diseases, particularly multiple sclerosis, using specified  
T polypeptides or related nucleic acid or ligand.  
T Claim 1; Page 167; 209pp; French.  
S  
X The present sequence represents a human protein, which is used in the  
C method of the invention. The specification describes a method which uses  
C at least one polypeptide or polynucleotide sequence belonging to the  
C perlecan, precursor of the retinol-binding plasma protein, precursor of  
C the ganglioside GM2 activator, calgranulin B or saposin B protein  
C families. The method is used for detecting, preventing or treating a  
C degenerative, neurological and/or auto-immune disease. The  
C polynucleotides and polypeptides are used for diagnosis, prognosis,  
C prevention and treatment of multiple sclerosis (in its various forms and  
C phases). They may also be useful in cases of e.g. Alzheimer's and  
C Parkinson's diseases, amyotrophic lateral sclerosis, rheumatoid  
C polyarthritis and lupus erythematosus, including use as vaccines and in  
C gene therapy (expression of sense or antisense sequences). They can also  
C be used to assess efficacy of potential therapeutic agents, particularly  
C compounds that reduce or inhibit toxicity towards glial cells  
X  
X Sequence 92 AA;  
Y Query Match 71.1%; Score 332; DB 4; Length 92;  
b Best Local Similarity 70.3%; Pred. No. 2.5e-30;  
Matches 64; Conservative 10; Mismatches 17; Indels 0; Gaps 0;  
Y 1 TKLEHLEGIINIFHOYSVRLGHYDTLTKRELKQLITKELNTLKNTKDQGTIDKIFQNL 60  
b 2 TKLEHLEGIINIFHOYSVRKGHFDLTLSKGLKQLITKELANTIKNIKDKRAVIDEIFQGL 61  
Y 61 DANQEQVSFKFVVLVDVLTITAHNDNIHKE 91  
b 62 DANQEQVDQFISLVAILKAAHYHTHKE 92  
ESUL7.13  
DA93649  
D ADA93649 standard; protein; 92 AA.  
X ADA93649;  
X 20-NOV-2003 (first entry)  
T Human calgranulin C protein SEQ ID NO:2.  
X inflammatory disease; calgranulin C; antiinflammatory; gene therapy;  
W vasculitis; Kawasaki disease; cystic fibrosis;  
W chronic inflammatory disease; ulcerative colitis; Crohn's disease;  
W chronic bronchitis; inflammatory arthritis; psoriatic arthritis;  
W rheumatoid arthritis; seronegative arthritis;  
W systemic onset juvenile rheumatoid arthritis; Sjogren's disease;  
W acute inflammation; human.  
X  
S Homo sapiens.  
X WO2003069341-A2.  
X 21-AUG-2003.  
D 17-FEB-2003; 2003WO-EP001575.  
X 15-FEB-2002; 2002US-00077600.  
X (SWIT-) SWITCH BIOTECH AG.  
A (SORG/) SORG C.  
A (ROTH/) ROTH J.  
X Sorg C, Roth J;  
I

XX WPI; 2003-671681/63.  
DR N-PSDB; ADA93648.  
XX  
PT Diagnosing, treating or preventing inflammatory diseases comprises  
PT determining the amount and/or concentration of CALGRANULIN C polypeptide  
PT and/or nucleic acids encoding the polypeptide present in a biological  
PT sample.  
XX  
XX Claim 7; Page 64; 64pp; English.  
PS The present invention describes a method for diagnosing inflammatory  
XX diseases, which comprises determining the amount and/or concentration of  
CC calgranulin C polypeptide and/or nucleic acids encoding the polypeptide  
CC present in the biological sample. Also described are methods for treating  
CC or preventing an inflammatory disease in a mammal, and medical treatment  
CC of the mammal, where the treatment is based on the stage of the disease  
CC to be treated or prevented. Calgranulin C has antiinflammatory activity  
CC and can be used in gene therapy. The method is useful for diagnosing,  
CC treating or preventing inflammatory diseases, e.g. vasculitis  
CC (particularly Kawasaki disease), cystic fibrosis, chronic inflammatory  
CC diseases like ulcerative colitis or Crohn's disease, rheumatoid arthritis,  
CC inflammatory arthritis (e.g. psoriatic arthritis), juvenile arthritis or  
CC seronegative arthritis), systemic onset juvenile rheumatoid arthritis  
CC (Sjogren's disease), acute inflammation above the background of a  
CC chronic inflammation, an acquired infection on the background of an  
CC inflammatory disease, or an exacerbation of an already present disease.  
CC The method is also useful for diagnosing specific stages of inflammatory  
CC diseases, for determining the risk of relapse, and for discriminating  
CC between diseases with similar symptoms. The present sequence represents  
CC human calgranulin C, which is used in the exemplification of the present  
CC invention.  
XX  
SQ Sequence 92 AA;  
Y Query Match 71.1%; Score 332; DB 7; Length 92;  
b Best Local Similarity 70.3%; Pred. No. 2.5e-30;  
Matches 64; Conservative 10; Mismatches 17; Indels 0; Gaps 0;  
Y 1 TKLEHLEGIINIFHOYSVRLGHYDTLTKRELKQLITKELNTLKNTKDQGTIDKIFQNL 60  
b 2 TKLEHLEGIINIFHOYSVRKGHFDLTLSKGLKQLITKELANTIKNIKDKRAVIDEIFQGL 61  
Y 61 DANQEQVSFKFVVLVDVLTITAHNDNIHKE 91  
b 62 DANQEQVDQFISLVAILKAAHYHTHKE 92  
RESULT 14  
ABG27582  
ID ABG27582 standard; protein; 95 AA.  
XX  
AC ABG27582;  
XX  
DT 18-FEB-2002 (first entry)  
XX  
DE Novel human diagnostic protein #27573.  
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US008631.  
XX  
PR 31-MAR-2000; 2000US-00540217.  
PR 23-AUG-2000; 2000US-00649167.  
XX  
PA (HYSE-) HYSEQ INC.

XX	Dmanac RT, Liu C, Tang YT;
XX	
PI	
XX	
DR	WPI; 2001-639362/73.
DR	N-PSDB; AAS91769.
XX	
XX	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity.
XX	
XX	Claim 20; SEQ ID NO 57941; 103pp; English.
PS	
XX	
CC	The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC	sequences. (I) is useful as hybridisation probes, polymerase chain
CC	reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC	and in recombinant production of (II). The polynucleotides are also used
CC	in diagnostics as expressed sequence tags for identifying expressed
CC	genes. (I) is useful in gene therapy techniques to restore normal
CC	activity of (II) or to treat disease states involving (II). (II) is
CC	useful for generating antibodies against it, detecting or quantitating a
CC	polypeptide in tissue, as molecular weight markers and as a food
CC	supplement. (II) and its binding partners are useful in medical imaging
CC	of sites expressing (II). (I) and (II) are useful for treating disorders
CC	involving aberrant protein expression or biological activity. The
CC	polypeptide and polynucleotide sequences have applications in
CC	diagnostics, forensics, gene mapping, identification of mutations
CC	responsible for genetic disorders or other traits to assess biodiversity
CC	and to produce other types of data and products dependent on DNA and
CC	amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC	amino acid sequences of the invention. Note: The sequence data for this
CC	patent did not appear in the printed specification, but was obtained in
CC	electronic format directly from WIPO at
CC	ftp.wipo.int/pub/published_pct_sequences
XX	
XX	Sequence 95 AA;
SQ	
	Query Match 65.5%; Score 306; DB 4; Length 95;
	Best Local Similarity 67.0%; Pred. No. 2.5e-27;
	Matches 61; Conservative 10; Mismatches 20; Indels 0; Gaps 0;
QY	1 TKLEHLEGITINTHQVSVRLGHYDTLTKBELKQLITKELPNTLKNTKDQGTIDKIFQNL 60
Db	:     :     :     :     :     :     :     :     :
	5 TKLEEHLLEGIVNISPKXSVRKGFHTLSKGELKQLTKELANTIKTKDKAVIDEIFQGL 64
QY	61 DANQDSQVSFKFVLVTVDYLITAHDNHKKE 91
Db	:     :     :     :     :     :     :     :
	65 DANQDSQVDFQEFISLVIAALKAAHYTHKE 95
RESULT 15	
AALW17062	
ID	AALW17062 standard; protein; 114 AA.
XX	
AC	AALW17062;
XX	
DT	16-JUL-1997 (first entry)
XX	
DE	Human multidrug resistance protein 14 (MRP14).
XX	
KW	Genetic engineering; MRP; multidrug resistance protein; transgenic;
KW	animal model; cell death inhibition; apoptosis; cell proliferation; HIV;
KW	human immunodeficiency virus; cancer; cystic fibrosis; neoplasia; tumour.
XX	
OS	Homo sapiens.
XX	
US	US5614397-A.
PN	
PD	25-MAR-1997.
XX	
XX	22-FEB-1994; 94US-00200016.
XX	
PR	22-FEB-1994; 94US-00200016.



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

M protein - protein search, using sw model

un on: March 5, 2004, 09:45:44 ; Search time 23 seconds  
(without alignments)  
204.259 Million cell updates/sec

itle: US-09-646-651C-1

erfect score: 467

equences: 1 TKLEDHLEGIINPHQYSVR.....EFVVLTVDTLTAHDNIHKE 91

coring table: BLOSUM62

earched: Gapop 10.0 , Gapext 0.5

total number of hits satisfying chosen parameters: 389414

inimum DB seq length: 0  
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

atabase : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Query Match	Length	DB ID	Description
1	467	100.0	91	3	US-08-794-000-2
2	371	79.4	92	2	US-08-568-310D-19
3	371	79.4	92	4	US-09-270-455-19
4	366	78.4	90	4	US-09-263-312-3
5	366	78.4	90	4	US-09-826-589-3
6	366	78.4	90	4	US-09-826-589-4
7	332	71.1	92	2	US-08-568-310D-20
8	332	71.1	92	4	US-09-270-455-20
9	235	50.3	51	2	US-08-568-310D-2
10	235	50.3	51	4	US-09-270-455-2
11	213.5	45.7	114	1	US-08-385-241-3
12	213.5	45.7	114	4	US-09-214-272-4
13	211.5	45.3	109	1	US-07-987-272A-8
14	206	44.1	50	4	US-09-263-312-2
15	206	44.1	50	4	US-09-826-589-2
16	204.5	43.8	113	2	US-08-918-727-7
17	204.5	43.8	113	3	US-09-205-680A-7
18	181	38.8	95	4	US-09-919-172-102
19	181	38.8	95	4	US-09-976-594-467
20	164	35.1	92	2	US-09-051-589-1
21	163	34.9	88	1	US-07-987-272A-1
22	163	34.9	89	1	US-07-987-272A-14
23	162	34.7	91	1	US-07-987-272A-11
24	162	34.7	92	2	US-08-918-727-5
25	162	34.7	92	3	US-09-205-680A-5
26	156	33.4	101	1	US-08-190-560-2
27	156	33.4	101	1	US-08-469-277-2

28 156 33.4 101 2 US-08-468-946-2 Sequence 2, Appli  
29 156 33.4 101 2 US-08-468-942-2 Sequence 2, Appli  
30 156 33.4 101 4 US-09-298-625-2 Sequence 2, Appli  
31 151 32.3 93 1 US-07-987-272A-7 Sequence 7, Appli  
32 151 32.3 93 1 US-07-987-272A-16 Sequence 16, Appli  
33 151 32.3 93 1 US-08-385-241-1 Sequence 1, Appli  
34 151 32.3 93 4 US-09-214-272-2 Sequence 6, Appli  
35 150 32.1 105 2 US-08-918-727-6 Sequence 6, Appli  
36 150 32.1 105 3 US-09-205-680A-6 Sequence 2, Appli  
37 148 31.7 97 1 US-07-662-198B-2 Sequence 3, Appli  
38 140.5 30.1 98 2 US-08-918-727-3 Sequence 3, Appli  
39 140.5 30.1 98 3 US-09-205-680A-3 Sequence 11, Appli  
40 140.5 30.1 98 3 US-09-048-889-11 Sequence 17, Ap  
41 140 30.0 92 4 US-09-621-976-7537 Sequence 17, Ap  
42 137 29.3 76 1 US-07-987-272A-17 Sequence 7524, Ap  
43 136 29.1 88 4 US-09-621-976-7524 Sequence 7526, Ap  
44 136 29.1 88 4 US-09-621-976-7526 Sequence 10, Appl  
45 136 29.1 89 1 US-07-987-272A-10

ALIGNMENTS

RESULT 1  
US-08-794-000-2  
; Sequence 2, Application US/08794000  
; Patent No. 6087123  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: Metal-Containing Ribonucleotide Polypeptides  
; NUMBER OF SEQUENCES: 4  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPC)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/794,000  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/DE96/01337  
; FILING DATE: 17-JUL-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE 195 25 992.0  
; FILING DATE: 17-JUL-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE 195 30 500.0  
; FILING DATE: 18-AUG-1995  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 91 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-794-000-2

Query Match 100.0%; Score 467; DB 3; Length 91;  
Best Local Similarity 100.0%; Pred. NO. 4.4e-49;  
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKLEDHLEGIINPHQYSVRLGHVDTLTKLEKOLITKELPNTLKNTKDQGTIDKIFONL 60  
DB 1 TKLEDHLEGIINPHQYSVRLGHVDTLTKLEKOLITKELPNTLKNTKDQGTIDKIFONL 60  
QY 61 DANQDEQVSFKSFVVLVTVTLTAHDNIHKE 91  
DB 61 DANQDEQVSFKSFVVLVTVTLTAHDNIHKE 91

RESULT 2  
US-08-568-310D-19  
; Sequence 19, Application US/08568310D

Patent No. 5976932  
GENERAL INFORMATION:  
APPLICANT: HITOMI, JIRO  
APPLICANT: YAMAGUCHI, KEN  
APPLICANT: YAMAMURA, TOKUJIRO  
APPLICANT: KIMURA, TATSUJI  
TITLE OF INVENTION: NOVEL CALCIUM-BINDING PROTEINS  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: WYATT GERBER, MELLER & O'ROURKE  
STREET: 99 PARK AVENUE  
STREET: 6th FLOOR  
CITY: NEW YORK CITY  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10016

COMPUTER READABLE FORM:  
MEDIUM TYPE: DISKETTE, 3.50 INCH, 720 KB  
MEDIUM TYPE: STORAGE  
COMPUTER: IBM-PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS 6.2  
SOFTWARE: WORDPERFECT 6.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/568,310D  
FILING DATE: DECEMBER 6, 1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 27101  
REFERENCE/DOCKET NUMBER: 3316  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)953-3350  
TELEFAX: (212)953-3352  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 92  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
PUBLICATION INFORMATION:  
RELEVANT RESIDUES IN SEQ ID NO: 19:  
RELEVANT RESIDUES IN SEQ ID NO: FROM 1 TO 92

US-08-568-310D-19

Query Match 79.4%; Score 371; DB 2; Length 92;  
Best Local Similarity 81.3%; Pred. No. 1.7e-37;  
Matches 74; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

QY 1 TKLEHLEGIINIFHOYSVRLGHVDTLTKRELKQLITKELPNTLNKTKDQGTIDKIFQNL 60  
DB 2 TKLEHLEGIINIFHOYSVRLGHVDTLTKRELKQLITKELPNTLNKTKDQGTIDKIFQDL 61

QY 61 DANQDEQVSPKFEFVVLVTDVLITAHDNHKE 91  
DB 62 DADKDGAVSFEFVVLVSRVLKTAHDIHKE 92

RESULT 3  
US-09-270-455-19  
Sequence 19, Application US/09270455  
Patent No. 6313267  
GENERAL INFORMATION:  
APPLICANT: HITOMI, JIRO  
APPLICANT: YAMAGUCHI, KEN  
APPLICANT: YAMAMURA, TOKUJIRO  
APPLICANT: KIMURA, TATSUJI  
TITLE OF INVENTION: NOVEL CALCIUM-BINDING PROTEINS  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:

QY 1 TKLEHLEGIINIFHOYSVRLGHVDTLTKRELKQLITKELPNTLNKTKDQGTIDKIFQNL 60  
DB 2 TKLEHLEGIINIFHOYSVRLGHVDTLTKRELKQLITKELPNTLNKTKDQGTIDKIFQDL 61

QY 61 DANQDEQVSPKFEFVVLVTDVLITAHDNHKE 91

DB 62 DADKDGAVSFEFVVLVSRVLKTAHDIHKE 92

US-09-270-455-19  
Sequence 19, Application US/09270455  
Patent No. 6313267  
GENERAL INFORMATION:  
APPLICANT: HITOMI, JIRO  
APPLICANT: YAMAGUCHI, KEN  
APPLICANT: YAMAMURA, TOKUJIRO  
APPLICANT: KIMURA, TATSUJI  
TITLE OF INVENTION: NOVEL CALCIUM-BINDING PROTEINS  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:

ADDRESSEE: WYATT GERBER, MELLER & O'ROURKE  
STREET: 99 PARK AVENUE  
STREET: 6th FLOOR  
CITY: NEW YORK CITY  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10016  
COMPUTER READABLE FORM:  
MEDIUM TYPE: DISKETTE, 3.50 INCH, 720 KB  
MEDIUM TYPE: STORAGE  
COMPUTER: IBM-PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS 6.2  
SOFTWARE: WORDPERFECT 6.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/270,455  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/568,310  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: KLEIN, MILTON  
REGISTRATION NUMBER: 27101  
REFERENCE/DOCKET NUMBER: 3316  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)953-3350  
TELEFAX: (212)953-3352  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 92  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
PUBLICATION INFORMATION:  
RELEVANT RESIDUES IN SEQ ID NO: 19: FROM 1 TO 92

US-09-270-455-19

Query Match 79.4%; Score 371; DB 4; Length 92;  
Best Local Similarity 81.3%; Pred. No. 1.7e-37;  
Matches 74; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

QY 1 TKLEHLEGIINIFHOYSVRLGHVDTLTKRELKQLITKELPNTLNKTKDQGTIDKIFQNL 60  
DB 2 TKLEHLEGIINIFHOYSVRLGHVDTLTKRELKQLITKELPNTLNKTKDQGTIDKIFQDL 61

QY 61 DANQDEQVSPKFEFVVLVTDVLITAHDNHKE 91  
DB 62 DADKDGAVSFEFVVLVSRVLKTAHDIHKE 92

RESULT 4  
US-09-263-312-3  
Sequence 3, Application US/09263312  
Patent No. 6555340  
GENERAL INFORMATION:  
APPLICANT: Schmidt, Ann Marie  
APPLICANT: Stern, David  
TITLE OF INVENTION: Extracellular No. 6555340el RAGE Binding Protein (EN-RAGE) and  
TITLE OF INVENTION: Uses Thereof  
FILE REFERENCE: 6575/55873-A  
CURRENT APPLICATION NUMBER: US/09/263,312  
CURRENT FILING DATE: 1999-03-05  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 3  
LENGTH: 90  
TYPE: PRT  
ORGANISM: Human  
US-09-263-312-3

Query Match 78.4%; Score 366; DB 4; Length 90;  
Best Local Similarity 81.1%; Pred. No. 6.7e-37;

QY 1 TKLEHLEGIINIFHOYSVRLGHVDTLTKRELKQLITKELPNTLNKTKDQGTIDKIFQNL 60  
DB 2 TKLEHLEGIINIFHOYSVRLGHVDTLTKRELKQLITKELPNTLNKTKDQGTIDKIFQDL 61

QY 61 DANQDEQVSPKFEFVVLVTDVLITAHDNHKE 91

DB 62 DADKDGAVSFEFVVLVSRVLKTAHDIHKE 92

US-09-263-312-3  
Sequence 3, Application US/09263312  
Patent No. 6555340  
GENERAL INFORMATION:  
APPLICANT: Schmidt, Ann Marie  
APPLICANT: Stern, David  
TITLE OF INVENTION: Extracellular No. 6555340el RAGE Binding Protein (EN-RAGE) and  
TITLE OF INVENTION: Uses Thereof  
FILE REFERENCE: 6575/55873-A  
CURRENT APPLICATION NUMBER: US/09/263,312  
CURRENT FILING DATE: 1999-03-05  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 3  
LENGTH: 90  
TYPE: PRT  
ORGANISM: Human  
US-09-263-312-3

Query Match 78.4%; Score 366; DB 4; Length 90;  
Best Local Similarity 81.1%; Pred. No. 6.7e-37;

QY 1 TKLEHLEGIINIFHOYSVRLGHVDTLTKRELKQLITKELPNTLNKTKDQGTIDKIFQNL 60  
DB 2 TKLEHLEGIINIFHOYSVRLGHVDTLTKRELKQLITKELPNTLNKTKDQGTIDKIFQDL 61

QY 61 DANQDEQVSPKFEFVVLVTDVLITAHDNHKE 91

DB 62 DADKDGAVSFEFVVLVSRVLKTAHDIHKE 92

US-09-263-312-3  
Sequence 3, Application US/09263312  
Patent No. 6555340  
GENERAL INFORMATION:  
APPLICANT: Schmidt, Ann Marie  
APPLICANT: Stern, David  
TITLE OF INVENTION: Extracellular No. 6555340el RAGE Binding Protein (EN-RAGE) and  
TITLE OF INVENTION: Uses Thereof  
FILE REFERENCE: 6575/55873-A  
CURRENT APPLICATION NUMBER: US/09/263,312  
CURRENT FILING DATE: 1999-03-05  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 3  
LENGTH: 90  
TYPE: PRT  
ORGANISM: Human  
US-09-263-312-3

Query Match 78.4%; Score 366; DB 4; Length 90;  
Best Local Similarity 81.1%; Pred. No. 6.7e-37;

Matches 73; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

Y 1 TKLEHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDOGTIDKIFQNL 60  
 b 1 TKLEHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDOGTIDKIFQNL 60  
 Y 61 DANODEQVSFKFVVLVDTVLITAHNDIHK 90  
 b 61 DADKDGAVSFEEFVVLVSRVLKTAHIDIHK 90

RESULT 5

S-09-826-589-3  
 Sequence 3, Application US/09826589  
 Patent No. 6670136  
 GENERAL INFORMATION:  
 APPLICANT: Schmidt, Ann Marie  
 APPLICANT: Stern, David  
 TITLE OF INVENTION: EXTRACELLULAR NOVEL RAGE BINDING PROTEIN (EN-RAGE) AND USES THERE  
 FILE REFERENCE: 0575/55873-B-PCT-US  
 CURRENT APPLICATION NUMBER: US/09/826,589  
 CURRENT FILING DATE: 2001-04-05  
 NUMBER OF SEQ ID NOS: 6  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 3  
 LENGTH: 90  
 TYPE: PRT  
 ORGANISM: Bovine  
 S-09-826-589-3

Query Match 78.4%; Score 366; DB 4; Length 90;  
 Best Local Similarity 81.1%; Pred. No. 6.7e-37;  
 Matches 73; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

Y 1 TKLEHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDOGTIDKIFQNL 60  
 b 1 TKLEHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDOGTIDKIFQNL 60  
 Y 61 DANODEQVSFKFVVLVDTVLITAHNDIHK 90  
 b 61 DADKDGAVSFEEFVVLVSRVLKTAHIDIHK 90

RESULT 6

S-09-826-589-4  
 Sequence 4, Application US/09826589  
 Patent No. 6670136  
 GENERAL INFORMATION:  
 APPLICANT: Schmidt, Ann Marie  
 APPLICANT: Stern, David  
 TITLE OF INVENTION: EXTRACELLULAR NOVEL RAGE BINDING PROTEIN (EN-RAGE) AND USES THERE  
 FILE REFERENCE: 0575/55873-B-PCT-US  
 CURRENT APPLICATION NUMBER: US/09/826,589  
 CURRENT FILING DATE: 2001-04-05  
 NUMBER OF SEQ ID NOS: 6  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 4  
 LENGTH: 90  
 TYPE: PRT  
 ORGANISM: Bovine  
 S-09-826-589-4

Query Match 78.4%; Score 366; DB 4; Length 90;  
 Best Local Similarity 81.1%; Pred. No. 6.7e-37;  
 Matches 73; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

Y 1 TKLEHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDOGTIDKIFQNL 60  
 b 1 TKLEHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDOGTIDKIFQNL 60  
 Y 61 DANODEQVSFKFVVLVDTVLITAHNDIHK 90  
 b 61 DADKDGAVSFEEFVVLVSRVLKTAHIDIHK 90

RESULT 7

US-08-568-310D-20  
 ; Sequence 20, Application US/08568310D  
 ; Patent No. 5976832  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HITOMI, JIRO  
 ; APPLICANT: YAMAGUCHI, KEN  
 ; APPLICANT: YAMAMURA, TOKUJIRO  
 ; APPLICANT: KIMURA, TATSUJI  
 ; TITLE OF INVENTION: NOVEL CALCIUM-BINDING PROTEINS  
 ; NUMBER OF SEQUENCES: 20  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: WYATT, GERBER, MELLER & O'ROURKE  
 ; STREET: 99 PARK AVENUE  
 ; STREET: 6th FLOOR  
 ; CITY: NEW YORK CITY  
 ; STATE: NEW YORK  
 ; COUNTRY: USA  
 ; ZIP: 10016  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: DISKETTE, 3.50 INCH, 720 Kb  
 ; MEDIUM TYPE: STORAGE  
 ; COMPUTER: IBM-PC COMPATIBLE  
 ; OPERATING SYSTEM: PC-DOS 6.2  
 ; SOFTWARE: WORDPERFECT 6.1  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/568,310D  
 ; FILING DATE: DECEMBER 6, 1995  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 7-70468 and 7-45564 (both Japan)  
 ; FILING DATE: 3/6/95 and 3/6/95, respectively  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: KLEIN, MILTON  
 ; REGISTRATION NUMBER: 27101  
 ; REFERENCE/DOCKET NUMBER: 3316  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212)953-3350  
 ; TELEFAX: (212)953-3352  
 ; INFORMATION FOR SEQ ID NO: 20:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 92  
 ; TYPE: amino acid  
 ; STRANDEDNESS:  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: CDNA  
 ; PUBLICATION INFORMATION:  
 ; RELEVANT RESIDUES IN SEQ ID NO: 20:  
 ; RELEVANT RESIDUES IN SEQ ID NO: FROM 1 TO 92  
 ; US-08-568-310D-20

Query Match 71.1%; Score 332; DB 2; Length 92;  
 Best Local Similarity 70.3%; Pred. No. 8.7e-33;  
 Matches 64; Conservative 10; Mismatches 17; Indels 0; Gaps 0;

QY 1 TKLEHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDOGTIDKIFQNL 60  
 Db 2 TKLEHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKDOGTIDKIFQNL 61  
 QY 61 DANODEQVSFKFVVLVDTVLITAHNDIHK 91  
 Db 62 DANODEQVDFEFLVLAALKAAHYHTHKE 92

RESULT 8

US-09-270-455-20  
 ; Sequence 20, Application US/09270455  
 ; Patent No. 6313267  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HITOMI, JIRO  
 ; APPLICANT: YAMAGUCHI, KEN

APPLICANT: YAMAMURA, TOKUJIRO  
 APPLICANT: KIMURA, TATSUUI  
 TITLE OF INVENTION: NOVEL CALCIUM-BINDING PROTEINS  
 NUMBER OF SEQUENCES: 20  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: WYATT, GERBER, MELLER & O'ROURKE  
 STREET: 99 PARK AVENUE  
 STREET: 6th FLOOR  
 CITY: NEW YORK CITY  
 STATE: NEW YORK  
 COUNTRY: USA  
 ZIP: 10016  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: DISKETTE, 3.50 INCH, 720 KB  
 MEDIUM TYPE: STORAGE  
 COMPUTER: IBM-PC COMPATIBLE  
 OPERATING SYSTEM: PC-DOS 6.2  
 SOFTWARE: WORDPERFECT 6.1  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/270,455  
 FILING DATE:  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/569,310  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: KLEIN, MILTON  
 REGISTRATION NUMBER: 27101  
 REFERENCE/DOCKET NUMBER: 3316  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212)953-3350  
 TELEFAX: (212)953-3352  
 INFORMATION FOR SEQ ID NO: 20:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 92  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 MOLECULE TYPE: CDNA  
 PUBLICATION INFORMATION:  
 RELEVANT RESIDUES IN SEQ ID NO: 20: FROM 1 TO 92  
 US-09-270-455-20

COUNTRY: USA  
 ZIP: 10016  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: DISKETTE, 3.50 INCH, 720 Kb  
 MEDIUM TYPE: STORAGE  
 COMPUTER: IBM-PC COMPATIBLE  
 OPERATING SYSTEM: PC-DOS 6.2  
 SOFTWARE: WORDPERFECT 6.1  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/568,310D  
 FILING DATE: 3/6/95 and 3/6/95, respectively  
 FILING DATE: 3/6/95 and 3/6/95, respectively  
 ATTORNEY/AGENT INFORMATION:  
 NAME: KLEIN, MILTON  
 REGISTRATION NUMBER: 27101  
 REFERENCE/DOCKET NUMBER: 3316  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212)953-3350  
 TELEFAX: (212)953-3352  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 51  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 PUBLICATION INFORMATION:  
 RELEVANT RESIDUES IN SEQ ID NO: 2:  
 RELEVANT RESIDUES IN SEQ ID NO: FROM 1 TO 51  
 US-08-568-310D-2

Query Match 50.3%; Score 235; DB 2; Length 51;  
 Best Local Similarity 90.0%; Pred.No. 2e-21;  
 Matches 45; Conservative 3; Mismatches 2; Indels 0;

QY 1 TKLEDHLEGIINIFHOYSVRLGHVDTLIKRELKOLITKELPNTLKNTKQD 50  
 Db 1 TKLEDHLEGIINIFHOYSVRLGHVDTLIKRELKOLITKELPNTLKNTKQD 50

RESULT 10  
 US-09-270-455-2  
 Sequence 2, Application US/09270455  
 Patent No. 6313267  
 GENERAL INFORMATION:  
 APPLICANT: HITOMI, JIRO  
 APPLICANT: YAMAGUCHI, KEN  
 APPLICANT: YAMAMURA, TOKUJRO  
 APPLICANT: KIMURA, TATSUJI  
 TITLE OF INVENTION: NOVEL CALCIUM-BINDING PROTEINS  
 NUMBER OF SEQUENCES: 20  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: WYATT, GERBER, MELLER & O'ROURKE  
 STREET: 99 PARK AVENUE  
 STREET: 6th FLOOR  
 CITY: NEW YORK CITY  
 STATE: NEW YORK  
 COUNTRY: USA  
 ZIP: 10016  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: DISKETTE, 3.50 INCH, 720 Kb  
 MEDIUM TYPE: STORAGE  
 COMPUTER: IBM-PC COMPATIBLE  
 OPERATING SYSTEM: PC-DOS 6.2  
 SOFTWARE: WORDPERFECT 6.1  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/270,455  
 FILING DATE:  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 PRIOR APPLICATION NUMBER: 08/568,310

SOFTWARE: PatentIn Release #1.0, Version #1.25

;;  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/987,272A  
;; FILING DATE: 05-MAR-1993  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: AU PK 2127  
;; FILING DATE: 05-FEB-1990  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: AU PK 4463  
;; FILING DATE: 05-SEP-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Brinkman, David W  
;; REGISTRATION NUMBER: 20,817  
;; REFERENCE/DOCKET NUMBER: DWB/1925/200259  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 202-861 3000  
;; TELEFAX: 202-822 0944  
;; TELEX: 6714627 CUSH  
;; INFORMATION FOR SEQ ID NO: 8:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 109 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; US-07-987-272A-8

Query Match 45.3%; Score 211.5; DB 1; Length 109;  
Best Local Similarity 45.1%; Pred. No. 3.7e-18;  
Matches 41; Conservative 25; Mismatches 24; Indels 1; Gaps 1;  
Qy 1 TKLEHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKQ 59  
Db 1 SLEARNIETIINTFHQYSVKLGHPDTLNQGEFKELVRKDLQNFLLKXKNELVIEHIMED 60

Qy 60 LDANQDEQVSKEFVLVTDVLITAHNTHK 90  
Db 61 LDTNADKQLSFEFTMLMARLTWASHEKMH 91

RESULT 14  
US-09-263-312-2  
;; Sequence 2, Application US/09263312  
;; Patent No. 655340  
;; GENERAL INFORMATION:  
;; APPLICANT: Schmidt, Ann Marie  
;; APPLICANT: Stern, David  
;; TITLE OF INVENTION: Extracellular No. 655340a1 RAGE Binding Protein (EN-RAGE) and  
;; TITLE OF INVENTION: Uses Thereof  
;; FILE REFERENCE: 0575/55873-A  
;; CURRENT APPLICATION NUMBER: US/09/263,312  
;; CURRENT FILING DATE: 1999-03-05  
;; NUMBER OF SEQ ID NOS: 5  
;; SOFTWARE: Patent In Ver. 2.1  
;; SEQ ID NO 2  
;; LENGTH: 50  
;; TYPE: PRT  
;; ORGANISM: Human  
;; FEATURE:  
;; NAME/KEY: UNSURE  
;; LOCATION: (47)  
;; OTHER INFORMATION: Xaa at this position is unknown  
;; US-09-263-312-2

Query Match 44.1%; Score 206; DB 4; Length 50;  
Best Local Similarity 82.0%; Pred. No. 6.3e-18;  
Matches 41; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 TKLEHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKQ 50  
Db 1 TKLEHLEGIINIGHQYSVRVGHFTLNKYLKQLGTKEPKTLQNXKQ 50

RESULT 15  
US-09-826-589-2  
;; Sequence 2, Application US/09826589  
;; Patent No. 6670136  
;; GENERAL INFORMATION:  
;; APPLICANT: Schmidt, Ann Marie  
;; APPLICANT: Stern, David  
;; TITLE OF INVENTION: EXTRACELLULAR NOVEL RAGE BINDING PROTEIN (EN-RAGE) AND USES THERE  
;; FILE REFERENCE: 0575/55873-B-PCT-US  
;; CURRENT APPLICATION NUMBER: US/09/826,589  
;; CURRENT FILING DATE: 2001-04-05  
;; NUMBER OF SEQ ID NOS: 6  
;; SOFTWARE: Patent In version 3.1  
;; SEQ ID NO 2  
;; LENGTH: 50  
;; TYPE: PRT  
;; ORGANISM: Human  
;; FEATURE:  
;; NAME/KEY: MISC\_FEATURE  
;; LOCATION: (47)..(47)  
;; OTHER INFORMATION: x=any amino acid  
;; US-09-826-589-2  
Query Match 44.1%; Score 206; DB 4; Length 50;  
Best Local Similarity 82.0%; Pred. No. 6.3e-18;  
Matches 41; Conservative 3; Mismatches 6; Indels 0; Gaps 0;  
Qy 1 TKLEHLEGIINIFHQYSVRLGHYDTLIKRELKQLITKELPNTLKNTKQ 50  
Db 1 TKLEHLEGIINIGHQYSVRVGHFTLNKYLKQLGTKEPKTLQNXKQ 50  
Search completed: March 5, 2004, 09:50:52  
Job time : 23 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.  
MM protein - protein search, using sw model  
Run on: March 5, 2004, 09:46:49 ; Search time 34 Seconds  
(without alignments)  
565.145 Million cell updates/sec

Title: US-09-646-651C-1  
Perfect score: 467  
Sequence: 1 TKLEDHLEGIINIFHOYSVR.....EFVVLVTDVLTIAHDNIHKE 91

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA.\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

result No.	Score	Query Match	Length	DB ID	Description
1	366	78.4	90	9	US-09-826-589-3
2	366	78.4	90	9	US-09-826-589-4
3	366	78.4	90	9	US-09-872-185B-11
4	366	78.4	90	9	US-09-872-185B-12
5	332	71.1	92	14	US-10-077-600-2
6	213.5	45.7	114	14	US-10-134-841-4
7	213.5	45.7	114	14	US-10-308-279-32
8	213.5	45.7	114	15	US-10-116-275-225
9	213.5	45.7	114	15	US-10-131-410-146
10	206	44.1	50	9	US-09-826-589-2
11	206	44.1	50	9	US-09-872-185B-9
12	204.5	43.8	112	14	US-10-205-219-161
13	204.5	43.8	113	10	US-09-492-026-7
14	192	41.1	46	9	US-09-864-761-41579
15	181	38.8	95	9	US-09-919-172-102

16 181 38.8 95 9 US-09-981-353-98  
17 178.5 38.2 113 14 US-10-134-841-3  
18 171 36.6 91 14 US-10-106-698-6907  
19 163 34.9 89 14 US-10-134-841-1  
20 163 34.9 101 9 US-09-393-433-2  
21 163 34.9 101 9 US-09-781-509-2  
22 163 34.9 101 14 US-10-269-643-2  
23 162 34.7 92 10 US-09-492-026-5  
24 162 34.7 92 10 US-09-919-039-184  
25 160 34.3 94 14 US-10-097-340-270  
26 159 34.0 89 14 US-10-316-253-46  
27 156 33.4 101 9 US-09-393-433-1  
28 156 33.4 101 9 US-09-781-509-1  
29 156 33.4 101 13 US-10-067-618-2  
30 156 33.4 101 13 US-10-135-152-2  
31 156 33.4 101 14 US-10-269-643-1  
32 154 33.0 93 15 US-10-094-886-58  
33 151 32.3 93 14 US-10-134-841-2  
34 151 32.3 477 15 US-10-161-927-62  
35 150 32.1 105 10 US-09-492-026-6  
36 150 32.1 105 14 US-10-097-340-272  
37 150 32.1 105 14 US-10-301-822-177  
38 150 32.1 134 9 US-09-925-302-694  
39 143 30.6 97 14 US-10-097-340-274  
40 143 30.6 97 14 US-10-171-311-206  
41 143 30.6 97 15 US-10-236-031B-40  
42 141 30.2 84 10 US-09-849-138-34  
43 140.5 30.1 98 10 US-09-492-026-3  
44 140.5 30.1 110 15 US-10-264-049-2820  
45 138 29.6 186 15 US-10-264-049-4104

## ALIGNMENTS

RESULT 1  
US-09-826-589-3  
; Sequence 3, Application US/09826589  
; Patent No. US20020106726A1  
; GENERAL INFORMATION:  
; APPLICANT: Schmidt, Ann Marie  
; TITLE OF INVENTION: EXTRACELLULAR NOVEL RAGE BINDING PROTEIN (EN-RAGE) AND USES THERE  
; FILE REFERENCE: 0575/55873-B-PCT-US  
; CURRENT APPLICATION NUMBER: US/09/826,589  
; CURRENT FILING DATE: 2001-04-05  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 90  
; TYPE: PRT  
; ORGANISM: Bovine  
US-09-826-589-3

Query Match 78.4%; Score 366; DB 9; Length 90;  
Best Local Similarity 81.1%; Pred. No. 1.2e-32;  
Matches 73; Conservative 9; Mismatches 8; Indels 0; Gaps 0;  
QY 1 TKLEDHLEGIINIFHOYSVRLGHVDLTILKRELKOLITKELPNTLKNYKQGTIDKIFQML 60  
DB 1 TKLEDHLEGIINIFHOYSVRVGHFDLTINKRELKOLITKELPNTLKNYKQGTIDKIFQDL 60  
QY 61 DANQDEQVSFKSFVVLVTDVLTIAHDNIHKE 90  
DB 61 DADKDGANSFEFVVLVSRVLKTAHIDIHK 90

RESULT 2  
US-09-826-589-4  
; Sequence 4, Application US/09826589  
; Patent No. US20020106726A1  
; GENERAL INFORMATION:  
; APPLICANT: Schmidt, Ann Marie

```
; APPLICANT: Stern, David
; TITLE OF INVENTION: EXTRACELLULAR NOVEL RAGE BINDING PROTEIN (EN-RAGE) AND USES THERE
; FILE REFERENCE: 0575/55873-B-PCT-US
; CURRENT APPLICATION NUMBER: US/09/826,589
; CURRENT FILING DATE: 2001-04-05
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Bovine
US-09-826-589-4

Query Match      78.4%; Score 366; DB 9; Length 90;
Best Local Similarity 81.1%; Pred. No. 1.2e-32;
Matches 73; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

QY 1 TKLEHLEGIINIFHOYSVRLGHYDTLKRELKQLITKELPNTLKNTKDOGTIDKIFQNL 60
   |||||
Db 1 TKLEHLEGIINIFHOYSVRLGHYDTLKRELKQLITKELPNTLKNTKDOGTIDKIFQNL 60
   |||||

QY 61 DANQDEQVSFKFVVLVTDVLITAHNDNIHK 90
   |||||
Db 61 DADKDGAVSFEEFVVLVSRVLTAKTAHDIHK 90
   |||||

RESULT 3
US-09-872-185B-11
; Sequence 11, Application US/09872185B
; Patent No. US20020122799A1
; GENERAL INFORMATION:
; APPLICANT: Stern, David M.
; APPLICANT: Herold, Kevin
; APPLICANT: Yan, Shi Du
; APPLICANT: Schmidt, Ann Marie
; APPLICANT: Lamster, Ira
; TITLE OF INVENTION: METHODS FOR TREATING INFLAMMATION
; FILE REFERENCE: 0575/64080
; CURRENT APPLICATION NUMBER: US/09/872,185B
; CURRENT FILING DATE: 2001-06-01
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Bovine
US-09-872-185B-11

Query Match      78.4%; Score 366; DB 9; Length 90;
Best Local Similarity 81.1%; Pred. No. 1.2e-32;
Matches 73; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

QY 1 TKLEHLEGIINIFHOYSVRLGHYDTLKRELKQLITKELPNTLKNTKDOGTIDKIFQNL 60
   |||||
Db 1 TKLEHLEGIINIFHOYSVRLGHYDTLKRELKQLITKELPNTLKNTKDOGTIDKIFQNL 60
   |||||

QY 61 DANQDEQVSFKFVVLVTDVLITAHNDNIHK 90
   |||||
Db 61 DADKDGAVSFEEFVVLVSRVLTAKTAHDIHK 90
   |||||

RESULT 4
US-09-872-185B-12
; Sequence 12, Application US/09872185B
; Patent No. US20020122799A1
; GENERAL INFORMATION:
; APPLICANT: Stern, David M.
; APPLICANT: Herold, Kevin
; APPLICANT: Yan, Shi Du
; APPLICANT: Schmidt, Ann Marie
; APPLICANT: Lamster, Ira
; TITLE OF INVENTION: METHODS FOR TREATING INFLAMMATION
; FILE REFERENCE: 0575/64080
```

```
; CURRENT APPLICATION NUMBER: US/09/872,185B
; CURRENT FILING DATE: 2001-06-01
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Bovine
US-09-872-185B-12
```

```
Query Match      78.4%; Score 366; DB 9; Length 90;
Best Local Similarity 81.1%; Pred. No. 1.2e-32;
Matches 73; Conservative 9; Mismatches 8; Indels 0; Gaps 0;
```

```
QY 1 TKLEHLEGIINIFHOYSVRLGHYDTLKRELKQLITKELPNTLKNTKDOGTIDKIFQNL 60
   |||||
Db 1 TKLEHLEGIINIFHOYSVRLGHYDTLKRELKQLITKELPNTLKNTKDOGTIDKIFQNL 60
   |||||

QY 61 DANQDEQVSFKFVVLVTDVLITAHNDNIHK 90
   |||||
Db 61 DADKDGAVSFEEFVVLVSRVLTAKTAHDIHK 90
   |||||
```

```
RESULT 5
US-10-077-600-2
; Sequence 2, Application US/10077600
; Publication No. US20030175713A1
; GENERAL INFORMATION:
; APPLICANT: Switch Biotech AG
; TITLE OF INVENTION: Method for diagnosis of inflammatory diseases using Calgranulin (
; FILE REFERENCE: S30274US
; CURRENT APPLICATION NUMBER: US/10/077,600
; CURRENT FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 92
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-077-600-2
```

```
Query Match      71.1%; Score 332; DB 14; Length 92;
Best Local Similarity 70.3%; Pred. No. 6.6e-29;
Matches 64; Conservative 10; Mismatches 17; Indels 0; Gaps 0;
```

```
QY 1 TKLEHLEGIINIFHOYSVRLGHYDTLKRELKQLITKELPNTLKNTKDOGTIDKIFQNL 60
   |||||
Db 2 TKLEHLEGIINIFHOYSVRLGHYDTLKRELKQLITKELPNTLKNTKDOGTIDKIFQNL 61
   |||||

QY 61 DANQDEQVSFKFVVLVTDVLITAHNDNIHK 91
   |||||
Db 62 DANQDEQVDFQEFISLVAIALAAAHVTHKE 92
   |||||
```

```
RESULT 6
US-10-134-841-4
; Sequence 4, Application US/10134841
; Publication No. US20030003482A1
; GENERAL INFORMATION:
; APPLICANT: HALLE, JORN-PETER
; APPLICANT: GOPPELT, ANDREAS
; TITLE OF INVENTION: MRP8/MRP14 heterodimer, or its
; TITLE OF INVENTION: individual components in combination, for treating and/or
; TITLE OF INVENTION: preventing skin diseases, wounds and/or wound-healing
; TITLE OF INVENTION: disturbances, having a reduced quantity of MRP8/MRP14
; TITLE OF INVENTION: heterodimers
; FILE REFERENCE: 50125/031002
; CURRENT APPLICATION NUMBER: US/10/134,841
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: US 60/322,925
; PRIOR FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: DE 10121254.2
; PRIOR FILING DATE: 2001-04-30
```







Matches 38; Conservative 24; Mismatches 28; Indels 1; Gaps 1;

y    1 TKLEHLEGIIIFHOYSVRLGHYDTTLIKRELKQLITKLPTLTKNTK-DQGTIDKIFON 59  
     :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:::  
b    7 SOLERSISLIIVFHQYSRKYGHPDFTLNKAFFKWKNQDLFNLAKREKENLLRDMED 66

y    60 LDANQEQQVSFKFEVVVLVTDLVIATAHDNIHK 90  
     |||||:|||:|||::~::  
b    67 LDTNQDNQLSFEECCMLMGKIIFACHEKLHE 97

RESULT 14

US-09-864-761-41579  
Sequence 41579, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:

APPLICANT: Penn, Sharron G.  
APPlicant: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR FILE REFERENCE: Aeo mica-X-1

CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29

NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annonax Sequence Listing Engine vers. 1.1  
SEQ ID NO 41579  
LENGTH: 46  
TYPE: prt  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC011666.18  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.  
MM protein - protein search, using sw model  
Run on: March 5, 2004, 09:43:38 ; Search time 20 Seconds  
(without alignments)  
437.671 Million cell updates/sec  
Title: US-09-646-651C-1  
Perfect score: 467  
Sequence: 1 TKLEHLEGIINFHOYSVR.....EFVVLVDVLIHAHNIHKE 91  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues  
Total number of hits satisfying chosen parameters: 283366  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries  
Database : PIR\_78: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	467	100.0	91	2 A55406	calgranulin c - pi
2	332	71.1	92	2 JC4712	S-100 calcium-bind
3	261.5	56.0	122	1 A42628	calgranulin B - bo
4	213.5	45.7	114	1 B31848	calgranulin B [val
5	204.5	43.8	113	1 JN0686	calgranulin B - ra
6	181	38.8	95	2 S24146	S-100 protein P -
7	178.5	38.2	113	1 S68242	calgranulin B - mo
8	167	35.8	95	1 S35985	S-100 protein alph
9	164	35.1	92	2 A26557	S-100 protein beta
10	163	34.9	89	1 I56163	calgranulin A - mo
11	163	34.9	91	1 BCBOIB	S-100 protein beta
12	163	34.9	92	2 A48015	S-100 protein beta
13	162	34.9	101	2 S06207	calvasculin - mous
14	162	34.7	92	1 BCHOIB	S-100 protein beta
15	161	34.5	94	1 BCBOIA	S-100 protein alph
16	160	34.3	94	1 BCHOIA	calvasculin - rat
17	160	34.3	101	2 S01759	calgranulin A - ra
18	159	34.0	89	1 JN0685	calgranulin A - huma
19	156	33.4	101	2 A48219	calgizzarin - huma
20	154	33.0	102	1 JQ1300	S-100 calcium-bind
21	153	32.8	100	2 A53217	placental calcium-
22	151	32.3	93	1 BCRUCF	calgranulin A [val
23	151	32.3	591	2 A48135	profilaggrin - hum
24	150	32.1	105	1 I37080	calgizzarin - huma
25	148	31.7	98	2 A41988	S-100 calcium-bind
26	147	31.5	306	2 A48118	major epidermal ca
27	144.5	30.9	79	1 KLPGI	calcium-binding pr
28	143.5	30.7	110	2 B48219	S-100 calcium-bind
29	140.5	30.1	98	2 JCS064	S-100 calcium-bind

30 139.5 29.9 79 1 JN0246 calcium-binding pr  
31 138 29.6 97 2 A30129 S-100 protein, lun  
32 136 29.1 90 1 BCHOY calyoclin - human  
33 135.5 29.0 97 2 A28489 calpactin I light  
34 135 28.9 79 1 KLBOI calcium-binding pr  
35 135 28.9 90 1 S27011 calyoclin - rabbit  
36 134.5 28.8 95 2 A31373 calpactin I light  
37 133.5 28.6 97 1 JH0663 calpactin I light  
38 133 28.5 99 2 S20342 calcium-binding pr  
39 131 28.1 89 2 A54314 calyoclin - mouse  
40 129.5 27.7 95 1 LUPG10 calpactin I light  
41 129.5 27.7 97 2 JCI139 calpactin I light  
42 129.5 27.7 97 2 B28489 calpactin I light  
43 128 27.4 90 2 B28363 calyoclin - rat  
44 127.5 27.3 217 2 JEO330 26-kDa Ca2+-bindin  
45 125.5 26.9 98 2 JCS065 calcium-binding pr

## ALIGNMENTS

## RESULT 1

A55406  
calgranulin c - pig  
C:Species: Sus scrofa domestica (domestic pig)  
C>Date: 10-Feb-1995 #sequence\_revision 10-Feb-1995 #text\_change 23-May-1997  
C:Accession: A55406  
R:DelliAngelica, E.C.; Schleicher, C.H.; Santome, J.A.  
J. Biol. Chem. 269, 28929-28936, 1994  
A:Title: Primary structure and binding properties of calgranulin C, a novel S100-like ca  
A:Reference number: A55406; MUID:95050708; PMID:7961855  
A:Accession: A55406  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-91 <DEL>  
A:Superfamily: S-100 protein; calmodulin repeat homology  
A:Keywords: calcium binding; EF hand  
F:48-80/Domain: calmodulin repeat homology <EF2>

Query Match 100.0%; Score 467; DB 2; Length 91;  
Best Local Similarity 100.0%; Pred. No. 1.6e-37;  
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKLEHLEGIINFHOYSVRLGHVDTLKELKOLITKELPNTLKNTKDKQGTIDKIFQNL 60  
Db 1 TKLEHLEGIINFHOYSVRLGHVDTLKELKOLITKELPNTLKNTKDKQGTIDKIFQNL 60  
QY 61 DANQDEQVSFKEFVVLVDVLIHAHNIHKE 91  
Db 61 DANQDEQVSFKEFVVLVDVLIHAHNIHKE 91

## RESULT 2

JC4712  
S-100 calcium-binding protein A12 - human  
N:Alternate names: calcium-binding amniotic fluid protein 1 (CAAF1); calgranulin C; calg  
utrophil protein  
C:Species: Homo sapiens (man)  
C>Date: 10-May-1996 #sequence\_revision 19-Jul-1996 #text\_change 13-Aug-1999  
C:Accession: JC4712; JC4717; JC4891; S56113; S56114  
R:Yamamura, T.; Hicomi, C.; Nagasaki, K.; Suzuki, M.; Takahashi, E.; Saito, S.; Tsukada,  
Biochem. Biophys. Res. Commun. 221, 356-360, 1996  
A:Title: Human CAAF1 gene - molecular cloning, gene structure, and chromosome mapping.  
A:Reference number: JC4712; MUID:96192053; PMID:8619860  
A:Accession: JC4712  
A:Molecule type: mRNA  
A:Residues: 1-92 <YAM>  
A:Cross-references: DDBJ:D83657; NID:g1502284; PIDN:BAA12030.1; PID:g1502285  
R:Marti, T.; Ertmann, K.D.; Gallin, M.Y.  
Biochem. Biophys. Res. Commun. 221, 454-458, 1996  
A:Title: Host-parasite interaction in human onchocerciasis: Identification and sequence  
A:Reference number: JC4717; MUID:96192069; PMID:8619876  
A:Accession: JC4717





```
C;Keywords: acetylated amino end; calcium binding; EF hand; heterodimer; inflammation;  
F;2-113/Product: calgranulin B #status predicted <MAT>  
F;11-45/Domain: calmodulin repeat homology <EF1>  
F;55-87/Domain: calmodulin repeat homology <EP2>  
F;2/Modified site: acetylated amino end (Ala) (in mature form) #status experimental  
E;60-91/Disulfide bonds: #status experimental  
F;103,105,107/Binding site: zinc (His) #status predicted  
F;107/Modified site: 3'-methylhistidine (His) #status experimental
```

Query Match      38.2%; Score 178.5; DB 1; Length 113;

Best Local Similarity    36.3%; Pred.No. 4.4e-10;

Matches    33; Conservative    27; Mismatches    30; Indels         1; Gaps           1;

```
QY          1   TKLEDHLEGIINIFHOYSVRLGHVDTLIKRELKOLITKELPNTLKNTX-DGGTIDKIQN 59  
             ::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|:  
DB          7   SQMERSITIILDTFOHSRKEGHPDLSKFQWVEAQLATPMKKERNEALINDIMED 66  
             :>:::~::~:|||||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
```

QY          60 LDANDEQVSFKFEFVLVTDLITAHDNHK 90  
             |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
DB          67 LDTNQNDQLSFEECMMLAKLIIPACHEKLHE 97  
             |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 8

S35985

S-100 protein alpha chain - weatherfish

C;Species: Misgurnus fossilis (weatherfish)

C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999

C;Accession: S35985

R:Ivanenkov, V.V.; Gerke, V.; Minin, A.A.; Plessmann, U.; Weber, K.  
*Mech. Dev.* 42, 151-158, 1993

A>Title: Transduction of Ca(2+) signals upon fertilization of eggs; identification of an

A;Reference number: S35985; UID:94031845; PMID:8217841

A;Accession: S35985

A:Molecule type: protein

A;Residues: 1-95 <IVA>

C;Superfamily: S-100 protein; calmodulin repeat homology

C;Keywords: calcium binding; EF hand

F;49-81/Domain: calmodulin repeat homology <BF2>

Query Match      35.8%; Score 167; DB 1; Length 95;

Best Local Similarity    40.7%; Pred.No. 4.4e-09;

Matches    35; Conservative    18; Mismatches    33; Indels         0; Gaps           0;

```
QY          1   TKLEDHLEGIINIFHOYSVRLGHVDTLIKRELKOLITKELPNTLKNTKDQGTTDKIFQM 60  
             ::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|::::|:  
DB          2   SGLEGAMESLIIKVHFHYSSKSGKYSLKAELSKLLGDELNDFLSASXDPPVWEKIMSGL 61  
             ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```

QY          61 DANQDEQVSFKFEFVLVTDLITAHHD 86  
             |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
DB          62 DENQDGVEDPQEFPVVVAALTACNE 87  
             |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 9

A26557

S-100 protein beta chain - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 25-Mar-1988 #sequence\_revision 04-Nov-1994 #text\_change 13-Aug-1999

C;Accession: A60046; S07357; A26557

R;Maeda, T.; Usui, H.; Araki, K.; Kuwano, R.; Takahashi, Y.; Suzuki, Y.  
*Nucleic Acids Res.* 12, 7455-7465, 1984

A>Title: Molecular cloning and the complete nucleotide sequence of cDNA for S-100

A;Reference number: A60046; UID:91359841; PMID:1653388

A;Accession: A60046

A>Status: translation not shown

A:Molecule type: DNA

A;Residues: 1-92 <MAE>

A;Cross-references: GB:S53527

R;Kuwno, R.; Usui, H.; Maeda, T.; Fukui, T.; Yamamari, N.; Ohtsuka, E.; Ikebara, M.; Ta-

A;Residues: 1-92 <KUW>  
A;Cross-references: ENBL:X01090; NID:G57174; PIDN:CAA25567.1; PID:G57175  
R;Dunn, R.; Landry, C.; O'Hanlon, D.; Dunn, J.; Allore, R.; Brown, I.; Marks, A.  
J. Biol. Chem. 262, 3562-3566, 1987  
A;Title: Reduction in S100 protein Beta-subunit mRNA in C6 rat glioma cells following t  
A;Reference number: A26557; MUID:87137648; PMID:3818655  
A;Accession: A26557  
A:Molecule type: mRNA  
A;Residues: 6-92 <DUN>  
A;Cross-references: GB:M15705  
C;Comment: S-100 protein occurs as alpha-beta heterodimers, alpha-alpha heterodimers, a  
C;Genetics:  
A;Introns: 46/3  
C;Superfamily: S-100 protein; calmodulin repeat homology  
C;Keywords: brain; calcium binding; dimer; EF hand; zinc  
F;2-92/Product: S-100 protein beta chain #status predicted <MAT>  
F;6-40/Domain: calmodulin repeat homology <EF1>

Query Match            35.1%; Score 164; DB 2; Length 92;  
Best Local Similarity   37.4%; Pred. No. 8.2e-09;  
Matches     34; Conservative   19; Mismatches   38; Indels       0; Gaps      0;

Qy    1   TKLEPHLEGIIINFIHQYSVRLGHYDTLIRKELKOLITKEPLNTLNKTDDQGTDIKIFONL 60  
        :||| : ||| : |::| ::||| : ||| : ||| : ||| :  
Db    2   SELEKANVALIIVFHQYSGREGDKHLKSSELKLNNLSHFLEEIKEQEVVDKWMTL 61

Qy    61 DANQEQQVSFKFEVLVTDLVIHTAHNIHKE 91  
        :||| : ||| : |::| ::||| : ||| : ||| : ||| :  
Db    62 DEGDGECDFOEFMAPFSVMVTTACHEFFEHE 92

RESULT 10  
IS6163  
calgranulin A - mouse  
N;Alternate names: calcium-binding protein MRP-8; CP-10 chemotactic protein; macrophage  
C;Species: Mus musculus (house mouse)  
C;Date: 26-Jul-1996 #sequence\_revision 23-May-1997 #text\_change 22-Jun-1999  
C;Accession: IS6163; A42488  
R;Lackmann, M.; Rajasekariah, P.; Ismaaa, S.E.; Jones, G.; Cornish, C.J.; Hu, S.; Simpso  
J. Immunol. 150, 2381-2391, 1993  
A;Title: Identification of a chemotactic domain of the pro-inflammatory S100 protein CP  
A;Reference number: IS6163; MUID:93203618; PMID:8454868  
A;Accession: IS6163  
A>Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A;Residues: 1-89 <LAC2>  
A;Cross-references: GB:S57123; NID:G298706; PIDN:AAB25840.1; PID:G298707  
R;Lackmann, M.; Cornish, C.J.; Simpson, R.J.; Moritz, R.L.; Geczy, C.L.  
J. Biol. Chem. 267, 7499-7504, 1992  
A;Title: Purification and structural analysis of a murine chemotactic cytokine (CP-10)  
A;Reference number: A42488; MUID:92218405; PMID:1559987  
A;Accession: A42488  
A:Molecule type: protein  
A;Residues: 2-77 <LAC2>  
A>Note: sequence extracted from NCBI backbone (NCBIP:94068)  
C;Genetics:  
A;Gene: MRP8  
C;Complex: homodimer; heterodimer and higher complexes with calgranulin B  
C;Superfamily: S-100 protein; calmodulin repeat homology  
C;Keywords: calcium binding; cytokine; EF hand; heterodimer; homodimer; inflammation  
F;2-89/Product: calgranulin A #status experimental <MAT>  
F;7-41/Domain: calmodulin repeat homology <BF1>  
F;46-78/Domain: calmodulin repeat homology <EF2>

Query Match            34.9%; Score 163; DB 1; Length 89;  
Best Local Similarity   33.0%; Pred. No. 9.8e-09;  
Matches     30; Conservative   25; Mismatches   32; Indels       4; Gaps      1;

Qy    1   TKLEPHLEGIIINFIHQYSVRLGHYDTLIRKELKOLITKEPLNTLNKTDDQGTDIKIFONL 60  
        :||| : ||| : |::| ::||| : ||| : ||| : ||| :  
Db    3   SELEKANSLIIDVHNYSIQGNHALYKNDFKKMVTTECFQVONI ----NIENLFREL 58

Qy    61 DANQEQQVSFKFEVLVTDLVIHTAHNIHKE 91

59 DINSNDNAINFELAMVIKGVASHKDSHKE 89

Db

RESULT 11

3CBOIB

S-100 protein beta chain - bovine  
C;Alternate names: neurocalcin delta-binding protein S100-beta  
C;Species: Bos primigenius taurus (cattle)  
C;Date: 31-May-1979 #sequence revision 14-Nov-1983 #text\_change 24-Nov-1999  
C;Accession: A91254; B91110; A90075; S54348; A03077  
R;Isobe, T.; Okuyama, T.  
R;J. Biochem. 89, 379-388, 1978  
A;Title: The amino-acid sequence of S-100 protein (PAP-I-b protein) and its relation to  
A;Reference number: A91254; MUID:79045265; PMID:710399  
A;Accession: A91254  
A;Molecule type: protein  
A;Residues: 'ESL', 5-91 <ISO>  
A;Experimental source: brain  
A;Note: this sequence has since been revised in reference A91110  
R;Isobe, T.; Okuyama, T.  
R;J. Biochem. 116, 79-86, 1981  
A;Title: The amino-acid sequence of the alpha subunit in bovine brain S-100 a protein.  
A;Reference number: A91110; MUID:81236562; PMID:7250124  
A;Accession: B91110  
A;Molecule type: protein  
A;Residues: 1-91 <IS2>  
R;Baudier, J.; Gerard, D.  
Biochemistry 22, 3360-3369, 1983  
A;Title: Ions binding to S100 proteins: structural changes induced by calcium and zinc  
A;Reference number: A90471; MUID:84000339; PMID:6615778  
A;Contents: annotation; metal ion-binding properties  
R;Marshak, D.R.; Umekawa, H.; Watterson, D.M.; Hidaka, H.  
Arch. Biochem. Biophys. 240, 777-780, 1985  
A;Title: Structural characterization of the calcium binding protein S100 from adipose ti  
A;Reference number: A90075; MUID:85278169; PMID:4026304  
A;Accession: A90075  
A;Molecule type: protein  
A;Residues: 1-91 <MAR>  
A;Experimental source: adipose tissue  
R;Okazaki, K.; Obata, N.H.; Inoue, S.; Hidaka, H.  
J. Biochem. J. 306, 551-555, 1995  
A;Title: S100-beta is a target protein of neurocalcin delta, an abundant isoform in glia  
A;Reference number: S54343; MUID:95194333; PMID:7887910  
A;Accession: S54348  
A;Molecule type: protein  
A;Residues: 56-61, 'N', 63-79, 'V' <OKA>  
R;Comment: The S-100 protein is composed of two related polypeptide chains, alpha and be  
rain proteins, S-100 is also found in a variety of other tissues.  
R;Comment: S-100 is an intracellular protein that weakly binds calcium. It binds zinc ve  
is with different affinities exist for both ions on each monomer. Physiological concentr  
ium-binding sites.  
R;Superfamily: S-100 protein; calmodulin repeat homology  
R;Keywords: blocked amino end; brain; calcium binding; EF hand; zinc  
R;5-39/Domain: calmodulin repeat homology <EF1>  
R;48-80/Domain: calmodulin repeat homology <EF2>  
R;1/Modified site: blocked amino end (Ser) (probably acetylated) #status experimental  
R;18,21,23,26,31/Binding site: calcium (Ser, Glu, Asp, Lys, Glu) #status predicted  
R;61,63,65,67,72/Binding site: calcium (Asp, Asp, Glu, Glu) #status predicted

Query Match 34.9%; Score 163; DB 1; Length 91;  
Best Local Similarity 36.3%; Pred. No. 1e-08;  
Matches 33; Conservative 20; Mismatches 38; Indels 0; Gaps 0;

1 TKLEDHLEGINIPHOYSVRLGHVDTLIKRELKQITKELPNTLNKTKDQGTIDKIFQNL 60

1 SELEKAVVALIDPHQYSGREGDKHLKSELKLNELSHFLSEIKESQEVVDKVMETL 60

61 DANQDEQVSKEFVVLVTDVLITAHDNHKE 91

61 DSDGDGECDFQEFMAFVAMITTACHEFFEHE 91

RESULT 12

A48015

S-100 protein beta chain - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 02-Jun-1995 #sequence\_revision 02-Jun-1995 #text\_change 24-Nov-1999  
C;Accession: A48015  
R;Jiang, H.; Shah, S.; Hilt, D.C.  
J. Biol. Chem. 268, 20502-20511, 1993  
A;Title: Organization, sequence, and expression of the murine S100beta gene. Transcripti  
A;Reference number: A48015; MUID:93388628; PMID:8376406  
A;Accession: A48015  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-92 <JTA>  
A;Cross-references: GB:L22144; NID:9404768; PIDN:AAA03075.1; PID:9404769  
C;Genetics:  
A;Introns: 46/3  
C;Superfamily: S-100 protein; calmodulin repeat homology  
C;Keywords: acetylated amino end; calcium binding; EF hand  
F;2-92/Product: S-100 protein beta chain #status predicted <MAT>  
F;6-40/Domain: calmodulin repeat homology <EF1>  
F;49-81/Domain: calmodulin repeat homology <EF2>  
F;2/Modified site: acetylated amino end (Ser) (in mature form) #status predicted  
F;20,22,24,27,32/Binding site: calcium (Gly, Glu, Asp, Lys, Glu) #status predicted  
F;62,64,66,68,70,73/Binding site: calcium (Asp, Asp, Glu, Asp, Glu) #status predict  
Query Match 34.9%; Score 163; DB 2; Length 92;  
Best Local Similarity 37.4%; Pred. No. 1e-08;  
Matches 34; Conservative 18; Mismatches 39; Indels 0; Gaps 0;  
1 TKLEDHLEGINIPHOYSVRLGHVDTLIKRELKQITKELPNTLNKTKDQGTIDKIFQNL 60  
2 SELEKAVVALIDPHQYSGREGDKHLKSELKLNELSHFLSEIKESQEVVDKVMETL 61  
61 DANQDEQVSKEFVVLVTDVLITAHDNHKE 91  
62 DSDGDGECDFQEFMAFVAMITTACHEFFEHE 92

RESULT 13

S06207

calvasculin - mouse  
C;Alternate names: calcium-binding protein mtsl; calcium-binding protein pEL98; placenta  
C;Species: Mus musculus (house mouse)  
C;Date: 31-Dec-1990 #sequence revision 31-Dec-1990 #text\_change 20-Jun-2000  
C;Accession: S06207; JH0097; S07981; A26803; A41411; I48674  
R;Ebraldize, A.; Tulchinsky, E.; Grigorian, M.; Afanasyeva, A.; Senin, V.; Revazova, E.;  
Genes Dev. 3, 1086-1093, 1989  
A;Title: Isolation and characterization of a gene specifically expressed in different me  
A;Reference number: S06207; MUID:89378739; PMID:2550322  
A;Accession: S06207  
A;Molecule type: mRNA  
A;Residues: 1-101 <EBR>  
A;Cross-references: EMBL:X16190; NID:954926; PIDN:CRA34316.1; PID:954927  
R;Tulchinsky, E.M.; Grigorian, M.S.; Ebraldize, A.K.; Malshina, N.I.; Lukanidin, E.M.  
Gene 87, 219-223, 1990  
A;Title: Structure of gene mtsl, transcribed in metastatic mouse tumor cells.  
A;Reference number: JH0097; MUID:90236313; PMID:2332170  
A;Accession: JH0097  
A;Molecule type: DNA  
A;Residues: 1-101 <TUL>  
A;Cross-references: GB:M36578; GB:M36579  
A;Experimental source: liver  
R;Tulchinsky, B.  
submitted to the EMBL Data Library, August 1989  
A;Reference number: S07981  
A;Accession: S07981  
A;Molecule type: DNA  
A;Residues: 1-47, 'VSGSXFNG', 56-57, 'RTDEAA', <TUL2>  
A;Cross-references: EMBL:X16094; NID:953249; PIDN:CRA34224.1; PID:953250  
R;Jackson-Grusby, L.L.; Swiergiel, J.; Linzer, D.I.H.  
Nucleic Acids Res. 15, 6677-6690, 1987



A:Title: A growth-related mRNA in cultured mouse cells encodes a placental calcium binding protein  
 A:Reference number: A26803; MUID:87316927; PMID:3628004  
 A:Accession: A26803  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-101 <JAC>  
 A:Cross-references: GB:J05935; NID:G50310; PIDN:CAA29282.1; PID:G50311  
 R:Goto, K.; Endo, H.; Fujiyoshi, T.  
 J. Biochem. 103, 48-53, 1988  
 A:Title: Cloning of the sequences expressed abundantly in established cell lines: identification of a novel protein  
 A:Reference number: A41411; MUID:98198109; PMID:3162911  
 A:Accession: A41411  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-101 <GCT>  
 A:Cross-references: GB:D00208; NID:G220569; PIDN:BAA00148.1; PID:G220570  
 R:Tulchinsky, E.; Kramerov, D.; Ford, H.L.; Reshetnyak, E.; Lukanidin, B.; Zain, S.  
 Oncogene 8, 79-86, 1993  
 A:Title: Characterization of a positive regulatory element in the mts1 gene.  
 A:Reference number: I48674; MUID:93141279; PMID:8423998  
 A:Accession: I48674  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-47, 'VSGSXFNQ', 48-54 <RES>  
 A:Cross-references: EMBL:X16094; NID:G53249; PIDN:CAA34224.1; PID:G53250  
 C:Comment: Gene mts1 is expressed in metastatic cells.  
 C:Genetics:  
 A:Gene: mts1  
 A:Introns: 47/3  
 A:Superfamily: S-100 protein; calmodulin repeat homology  
 C:Keywords: calcium binding; cancer; EF hand  
 F:7-41/Domain: calmodulin repeat homology <EF1>  
 F:50-82/Domain: calmodulin repeat homology <EF2>  
 Query Match 34.9%; Score 163; DB 2; Length 101;  
 Best Local Similarity 36.9%; Pred. No. 1.1e-08;  
 Matches 31; Conservative 23; Mismatches 30; Indels 0; Gaps 0;  
 QY 3 LEDHLEGIINIFHQYVRLGHVDTLTKRELKQLITKELNTLKNKDGTTIDKIFQNLDA 62  
 DB 5 LEEALDVIVSTFKYSGEGDKFKLNKTELKELLTRELPSFLGRKRTDEAAPQKVMNSLDS 64  
 QY 63 NDEQVSRKFEFVVLVDVLTITAH 86  
 DB 65 NRDNEVDFOEYCVFLSCIAMMCNE 88  
 RESULT 14  
 BCHUIA  
 S-100 protein beta chain [validated] - human  
 N:Alternate names: neural S-100 calcium-binding protein beta  
 C:Species: Homo sapiens (man)  
 C:Date: 04-Dec-1986 #sequence revision 06-Jan-1995 #text\_change 08-Dec-2000  
 C:Accession: A38364; A92972; A03076  
 R:Allore, R.J.; Friend, W.C.; O'Hanlon, D.; Neilson, K.M.; Baumal, R.; Dunn, R.J.; Markesbery, W.R.; Friend, W.C.; O'Hanlon, D.; Neilson, K.M.; Baumal, R.; Dunn, R.J.; Markesbery, W.R.  
 J. Biol. Chem. 265, 15537-15543, 1990  
 A:Title: Cloning and expression of the human S100beta gene.  
 A:Reference number: A38364; MUID:90368757; PMID:2394738  
 A:Accession: A38364  
 A:Molecule type: DNA  
 A:Residues: 1-92 <ALL>  
 A:Cross-references: GB:J05600; GB:M59486; NID:G337726; GB:M59487; GB:M59488  
 R:Jensen, R.; Marshak, D.R.; Anderson, C.; Lukas, T.J.; Watterson, D.M.  
 J. Neurochem. 45, 700-705, 1985  
 A:Title: Characterization of human brain S100 protein fraction: amino acid sequence of S100beta  
 A:Reference number: A92972; MUID:85291729; PMID:4031854  
 A:Accession: A92972  
 A:Molecule type: protein  
 A:Residues: 2-92 <JEN>  
 R:Baudier, J.; Glaser, N.; Haglid, K.; Gerard, D.  
 Biochim. Biophys. Acta 790, 164-173, 1984  
 A:Title: Purification, characterization and ion binding properties of human brain S100b  
 A:Reference number: A90653; MUID:85023393; PMID:6487634

A:Contents: annotation; metal ion-binding properties  
 C:Comment: This protein binds p53, tubulin and many other proteins at physiological concentrations  
 C:Comment: S-100 is an intracellular protein that binds calcium. It binds zinc more tightly than calcium. Different affinities exist for both ions on each monomer. Physiological concentrations of calcium and zinc are 10<sup>-6</sup> M and 10<sup>-7</sup> M, respectively.  
 C:Comment: This protein is expressed predominantly in brain tissue by astroglial cells.  
 C:Comment: The homodimer contains disulfide bonds, but the bond pattern has not been determined.  
 C:Genetics:  
 A:Gene: GDB:S100B  
 A:Cross-references: GDB:120360; OMIM:176990  
 A:Map position: 21q22.3-21q22.3  
 A:Introns: 46/3  
 A:Note: the first intron occurs before the initiator codon  
 C:Complex: homodimer; heterodimer with S-100 protein alpha chain (see PIR:BCHUIA)  
 C:Superfamily: S-100 protein; calmodulin repeat homology  
 C:Keywords: blocked amino end; brain; calcium binding; EF hand; heterodimer; homodimer; F2-92/Product: S-100 protein beta chain #status experimental <MAT>  
 F:6-40/Domain: calmodulin repeat homology <EF1>  
 F:49-81/Domain: calmodulin repeat homology <EF2>  
 F:2/Modified site: blocked amino end (Ser) (in mature form) (probably acetylated) #status predicted  
 F:19,22,24,27,32/Binding site: calcium (Ser, Glu, Asp, Lys, Glu) #status predicted  
 F:62,64,66,68,73/Binding site: calcium (Asp, Asp, Glu, Glu) #status predicted  
 Query Match 34.7%; Score 162; DB 1; Length 92;  
 Best Local Similarity 37.4%; Pred. No. 1.3e-08;  
 Matches 34; Conservative 18; Mismatches 39; Indels 0; Gaps 0;  
 QY 1 TKLEHLEGIINIFHQYVRLGHVDTLTKRELKQLITKELNTLKNKDGTTIDKIFQNL 60  
 DB 2 SELEKAMVALIDVFHQYSGREGDKHKLKSGELINNELSHFLKEIQEVVDVVMETL 61  
 QY 61 DANODEQVSRKFEFVVLVDVLTITAHNDIHKE 91  
 DB 62 DNDGDGECDFQEFMAFVAMVTTCHEFFEHE 92  
 RESULT 15  
 BCHUIA  
 S-100 protein alpha chain - bovine  
 C:Species: Bos primigenius taurus (cattle)  
 C:Date: 14-Nov-1983 #sequence revision 06-Feb-1995 #text\_change 24-Nov-1999  
 C:Accession: A24156; A91110; S54346; A03078  
 R:Kuwanon, R.; Maeda, T.; Usui, H.; Araki, K.; Yamakuni, T.; Ohshima, Y.; Kurihara, T.; FES Lett. 202, 97-101, 1986  
 A:Title: Molecular cloning of cDNA of S100alpha subunit mRNA.  
 A:Reference number: A24156; MUID:86248083; PMID:3755105  
 A:Accession: A24156  
 A:Molecule type: mRNA  
 A:Residues: 1-94 <KUW>  
 R:Isobe, T.; Okuyama, T.  
 Eur. J. Biochem. 116, 79-86, 1981  
 A:Title: The amino-acid sequence of the alpha subunit in bovine brain S-100 a protein.  
 A:Reference number: A91110; MUID:81236562; PMID:7250124  
 A:Accession: A91110  
 A:Molecule type: protein  
 A:Residues: 2-64, D', 66-94 <ISO>  
 R:Baudier, J.; Gerard, D.  
 Biochemistry 22, 3360-3369, 1983  
 A:Title: Ions binding to S100 proteins: structural changes induced by calcium and zinc  
 A:Reference number: A90471; MUID:84000339; PMID:6615778  
 A:Contents: annotation; metal ion-binding properties  
 R:Okazaki, K.; Obata, N.H.; Inoue, S.; Hidaka, H.  
 Biochem. J. 306, 551-555, 1995  
 A:Title: S100-beta is a target protein of neurocalcin delta, an abundant isoform in glioma  
 A:Reference number: S54343; MUID:95194333; PMID:7887910  
 A:Accession: S54346  
 A:Molecule type: protein  
 A:Residues: 24-33 <OKA>  
 C:Comment: The S-100 protein is composed of two related polypeptide chains, alpha and beta. Brain proteins, S-100 is also found in a variety of other tissues.  
 C:Comment: S-100 is an intracellular protein that weakly binds calcium. It binds zinc very weakly. With different affinities, exist for both ions on each monomer. Physiological concentrations of calcium and zinc are 10<sup>-6</sup> M and 10<sup>-7</sup> M, respectively.



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

M protein - protein search, using sw model

run on: March 5, 2004, 09:36:18 ; Search time 14 Seconds  
(without alignments)  
338.456 Million cell updates/sec

title: US-09-646-651C-1

effect score: 467

sequence: 1 TKLEDHLEGIINIFHOYSVR.....EFVVLVTDVLITAHNIHKE 91

scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

searched: 141681 seqs, 52070155 residues

total number of hits satisfying chosen parameters: 141681

minimum DB seq length: 0

maximum DB seq length: 2000000000

post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Query Match	Length	ID	Description
1	467	100.0	91	1 S112_PIG	P80310 sus scrofa
2	371	79.4	91	1 S112_BOVIN	P79105 bos taurus
3	332	71.1	91	1 S112_HUMAN	P80511 homo sapien
4	302	64.7	81	1 S112_RABIT	O77791 oryctolagus
5	263.5	56.4	122	1 S109_BOVIN	P28783 bos taurus
6	220	47.1	119	1 M126_CHICK	P28318 gallus gall
7	216.5	46.4	118	1 S109_RABIT	P50117 oryctolagus
8	213.5	45.7	114	1 S109_HUMAN	P06702 homo sapien
9	204.5	43.8	112	1 S109_RAT	P50116 rattus norv
10	181	38.8	95	1 S10F_HUMAN	P25815 homo sapien
11	178.5	38.2	112	1 S109_MOUSE	P31725 mus musculus
12	166	35.5	92	1 S10I_ICTUP	O91061 ictalurus p
13	164	35.1	91	1 S10B_RAT	P04631 rattus norv
14	164	34.9	98	1 S10Z_HUMAN	Q6wx98 homo sapien
15	163	34.9	98	1 S10B_MOUSE	P27005 mus musculus
16	163	34.9	88	1 S10B_RAT	P50115 rattus norv
17	163	34.9	91	1 S10B_BOVIN	P02638 bos taurus
18	163	34.9	91	1 S10B_MOUSE	P50114 mus musculus
19	163	34.9	93	1 S10A_RAT	P35467 rattus norv
20	163	34.9	101	1 S104_MOUSE	P07091 mus musculus
21	162	34.7	91	1 S10B_HUMAN	P04271 homo sapien
22	162	34.7	93	1 S10A_MOUSE	P56565 mus musculus
23	161	34.5	93	1 S10A_BOVIN	P02639 bos taurus
24	160	34.3	93	1 S10A_HUMAN	P23297 homo sapien
25	160	34.3	101	1 S10A_RAT	P05942 rattus norv
26	156	33.4	101	1 S104_HUMAN	P26447 homo sapien
27	154	33.0	102	1 S111_RABIT	P24480 oryctolagus
28	153	32.8	100	1 S104_BOVIN	P35466 bos taurus
29	151	32.3	93	1 S10B_HUMAN	P05109 homo sapien
30	150	32.1	105	1 S111_HUMAN	P31949 homo sapien
31	147	31.5	93	1 S10S_MOUSE	O88945 mus musculus
32	147	31.5	101	1 S111_CHICK	P24479 gallus gall
33	146	31.3	98	1 S111_MOUSE	P50543 mus musculus

ALIGNMENTS

RESULT 1

S112\_PIG

ID S112\_PIG STANDARD; PRT; 91 AA.

AC P80310; 78 1 S10D\_PIG

DT 01-FEB-1994 (Rel. 28, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Calgranulin C (CAGC).

GN S100A12.

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

OX NCBI\_TaxID=9823;

RN [1]

SEQUENCE.

RC TISSUE=Granulocyte;

RA MEDLINE=95050708; PubMed=7961855;

RA Dell'Angelica B.C., Schleicher C.H., Santone J.A.;

RT "Primary structure and binding properties of calgranulin C, a novel

RT S100-like calcium-binding protein from pig granulocytes.";

RL J. Biol. Chem. 269:28929-28936(1994).

CC -!- TISSUE SPECIFICITY: Found essentially in granulocytes with small

CC amounts found in lymphocytes.

CC -!- MISCELLANEOUS: In the absence of zinc binds one calcium ion per

CC molecule, in the presence of zinc binds two calcium ions per

CC molecule.

CC -!- SIMILARITY: Belongs to the S-100 family.

CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.

DR PIR; A55406; A55406.

DR HSSP; P80511; 1E8A.

DR InterPro; IPR001751; CaBP\_S100.

DR InterPro; IPR002048; EF-hand.

DR Pfam; PF00036; efhand; 1.

DR Pfam; PF01023; S\_100; 1.

DR ProDom; PD003407; CaBP\_S100; 1.

DR ProDom; PD000012; EF-hand; 1.

DR PROSITE; PS00018; EF\_HAND; FALSE NEG.

DR PROSITE; PS00303; S100\_CABP; 1.

KW Calcium-binding; Zinc; Metal-binding.

FT CA\_BIND 18 31 EF-HAND 1 (LOW AFFINITY) (BY SIMILARITY).

FT CA\_BIND 61 72 EF-HAND 2 (HIGH AFFINITY) (BY

FT SIMILARITY).

SQ SEQUENCE 91 AA; 10614 MW; B4204461432D7FCE CRC64;

Query Match 100.0%; Score 467; DB 1; Length 91;

Best Local Similarity 100.0%; Pred. No. 9.5e-37;

Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKLEDHLEGIINIFHOYSVRLGHVDTLTKELKQLITKELPNTLKNTKDQGTIDKIFQNL 60

DB 1 TKLEDHLEGIINIFHOYSVRLGHVDTLTKELKQLITKELPNTLKNTKDQGTIDKIFQNL 60

QY 61 DANQDEQVSFKFEFVLVTDVLITAHNIHKE 91

DB 61 DANQDEQVSFKFEFVLVTDVLITAHNIHKE 91

```
RESULT 2
S112_BOVIN STANDARD; PRT; 91 AA.
ID S112_BOVIN
AC P79105;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Calgranulin C (CAGC) (Calcium-binding protein in amniotic fluid 1)
DE (CAAF1) (RAGE binding protein).
GN S100A12 OR CAAF1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OC NCBI_TaxID=9913;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=Oesophagus; PubMed=8718672;
RX MEDLINE=96298783;
RA Hitomi J., Yamaguchi K., Kikuchi Y., Kimura T., Maruyama K.,
RA Nagasaki K.;
RA "A novel calcium-binding protein in amniotic fluid, CAAF1: its
RT molecular cloning and tissue distribution.";
RL J. Cell Sci. 109:805-815(1996).
[2]
RN SEQUENCE FROM N.A.
RP TISSUE=Lung;
RX MEDLINE=99325504; PubMed=1039917;
RA Hofmann M.A., Drury S., Fu C., Qu W., Taguchi A., Lu Y., Avila C.,
RA Kamtham N., Bierhaus A., Nawroth P., Neurath M.F., Slaterry T.,
RA Beach D., McClary J., Nagashima M., Morser J., Stern D.,
RA Schmidt A.M.;
RA "RAGE mediates a novel proinflammatory axis: a central cell surface
RT receptor for S100/calgranulin polypeptides.";
RL Cell 97:889-901(1999).
[3]
RN SEQUENCE FROM N.A.
RP TISSUE=Neutrophils;
RX MEDLINE=96192069; PubMed=8619876;
RA Marti T., Ertmann K.D., Gallin M.Y.;
RA "Host-parasite interaction in human onchocerciasis: identification
RT and sequence analysis of a novel human calgranulin.";
RL Biochem. Biophys. Res. Commun. 221:454-458(1996).
[4]
RN SEQUENCE FROM N.A.
RP TISSUE=Neutrophils;
RX MEDLINE=96332419; PubMed=8769108;
RA Ilg E.C., Troxler H., Buerigler D.M., Kuster T., Markert M.,
RA Guignard F., Hunziker P., Birchler N., Heizmann C.W.;
RA "Amino acid sequence determination of human S100A12 (P6, calgranulin
RT C, CGRP, CAAF1) by tandem mass spectrometry.";
RL Biochem. Biophys. Res. Commun. 225:146-150(1996).
[5]
RN SEQUENCE OF 1-20.
RX MEDLINE=95351965; PubMed=7626002;
RA Guignard F., Maue J., Markert M.;
RA "Identification and characterization of a novel human neutrophil
RT protein related to the S100 family.";
RL Biochem. J. 309:395-401(1995).
[6]
RN SEQUENCE OF 77-91, ANTIMICROBIAL ACTIVITY, AND MASS SPECTROMETRY.
RX TISSUE=Nasal mucus;
RX MEDLINE=21413725; PubMed=11522886;
RA Cole A.M., Kim Y.-H., Tahk S., Hong T., Weis P., Waring A.J., Ganz T.;
RA "Calcitermin, a novel antimicrobial peptide isolated from human airway
RT secretions.";
RL FEBS Lett. 504:5-10(2001).
[7]
RN X-RAY CRYSTALLOGRAPHY (1.95 ANGSTROMS).
RX MEDLINE=21065388; PubMed=11134923;
RA Moroz O.V., Antson A.A., Murshudov G.N., Maitland N.J., Dodson G.G.,
RA Wilson K.S., Skibshoj I., Lukyanidin E.M., Bronstein I.B.;
RT "The three-dimensional structure of human S100A12.";
RL Acta Crystallogr. D 57:20-29(2001).
```

Query Match 79.4%; Score 371; DB 1; Length 91;  
Best Local Similarity 81.3%; Pred. No. 7.1e-28;  
Matches 74; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

```
QY 1 TKLEHLEGINIFHOYSVRLGHYDTLTKRELKQLITKELPNTLTKNTKQGTIDKIFONL 60
DB 1 TKLEHLEGINIFHOYSVRLGHYDTLTKRELKQLITKELPNTLTKNTKQGTIDKIFQDL 60
```





"Rabbit polymorphonuclear neutrophils form 35S-labeled S-sulfo-calgranulin C when incubated with inorganic [35S]sulfate.";  
J. Biol. Chem. 271:19802-19809(1996).  
[2]  
SEQUENCE OF 45-82 FROM N.A.  
RC STRAIN=New Zealand white;  
MEDLINE=94198229; PubMed=8148323;  
XA Mori S., Goto K., Goto F., Mutakami K., Ohkawara S., Yoshinaga M.;  
"Dynamic changes in mRNA expression of neutrophils during the course  
of acute inflammation in rabbits.";  
Int. Immunol. 6:149-156(1994).  
-!- SIMILARITY: Belongs to the S-100 family.  
-!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; AF091849; AAC61771.1; -;  
EMBL; D17404; BAA04227.1; -;  
PIR; I46861; I46861.  
HSP; P80511; 1E8A.  
InterPro; IPR001751; CaBP\_S100.  
InterPro; IPR002048; EF-hand.  
Pfam; PF00036; efhand; 1.  
Pfam; PF01023; S\_100; 1.  
ProDom; PD003407; CaBP\_S100; 1.  
ProDom; PD000012; EF-hand; 1.  
PROSITE; PS00018; EF HAND; 1.  
PROSITE; PS00303; S100 CaBP; 1.  
Calcium-binding; Repeat.  
NON\_TER 1  
CA\_BIND 9 22 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).  
CA\_BIND 53 64 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).  
DOMAIN 103 118 2 X 8 AA TANDEM REPEATS OF G-H-G-H-  
G-H-S-H.  
REPEAT 103 110 1.  
REPEAT 111 118 2.  
SEQUENCE 118 AA; 13292 MW; 749611821AD5C41 CRC64;  
Query Match 46.4%; Score 216.5; DB 1; Length 118;  
Best Local Similarity 51.2%; Pred.No. 1.8e-13;  
Matches 42; Conservative 20; Mismatches 19; Indels 1; Gaps 1;  
10 IINIPHOYSVRLGHYDTLIKRELKQLITKELPNTL-KNTKDGQTIDKIFQNLDAHQEQV 68  
1 IINIPHOYSVGVPRDLSQKFKQLVQKELHFLKKEARDEKAINDIMEDLDTNQKQL 60  
69 SFKEFVLVTDVLIATADNIHK 90  
61 SFEEFVILMARLVHASHEEMHK 82  
RESULT 8  
1109 HUMAN STANDARD; PRT; 114 AA.  
D S109 HUMAN Q66702; Q66702; Q66702; Q66702; Q66702; Q66702; Q66702; Q66702; Q66702; Q66702;  
01-JAN-1988 (Rel. 06, Created)  
01-JAN-1988 (Rel. 06, Last sequence update)  
10-OCT-2003 (Rel. 42, Last annotation update)  
Calgranulin B (Migration inhibitory factor-related protein 14)  
(MRP-14) (P14) (Leukocyte LI complex heavy chain) (S100 calcium-  
binding protein A9) (Calprotectin L1H subunit).  
S100A9 OR MRP14 OR CAGB.  
Homo sapiens (Human).  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
NCBI\_TaxID=9606;  
[1]

SEQUENCE FROM N.A.  
RX MEDLINE=8039059; PubMed=3313057;  
RA Odink K., Gerletti N., Bruggen J., Clerc R.G., Tarcsay L., Zwaldo G.,  
Gerhards G., Schlegel R., Sorg C.;  
RT "Two calcium-binding proteins in infiltrate macrophages of rheumatoid  
arthritis.";  
RL Nature 330:80-82(1987).  
[2]  
SEQUENCE FROM N.A.  
RX MEDLINE=88302148; PubMed=3405210;  
RA Lagasse E., Clerc R.G.;  
RT "Cloning and expression of two human genes encoding calcium-binding  
proteins that are regulated during myeloid differentiation.";  
RL Mol. Cell. Biol. 8:2402-2410(1988).  
[3]  
SEQUENCE FROM N.A.  
RX MEDLINE=89255276; PubMed=2656677;  
RA Murao S., Collart F.R., Huberman E.;  
RT "A protein containing the cystic fibrosis antigen is an inhibitor of  
protein kinases.";  
RL J. Biol. Chem. 264:8356-8360(1989).  
[4]  
SEQUENCE FROM N.A. AND VARIANT ARG-20.  
RX Wang M., Xu X., Cai Y., Xu H., Han Y., Xu Z., Wu M.;  
RA "Human gene for migration inhibitory factor-related protein 14  
(MRP14), variant allele.";  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
[5]  
SEQUENCE FROM N.A.  
RC TISSUE=lung;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Klausner R.D., Feingold E.A., Grouse L.H., Derge J.G.,  
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,  
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,  
Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,  
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
[6]  
SEQUENCE OF 84-114, AND PHOSPHORYLATION.  
RX MEDLINE=90044075; PubMed=2478889;  
RA Edgeworth J., Freemont P., Hogg N.;  
RT "Ionomycin-regulated phosphorylation of the myeloid calcium-binding  
protein p14.";  
RL Nature 342:189-192(1989).  
[7]  
SEQUENCE OF 11-19; 26-37 AND 94-107.  
RC TISSUE=Keratinocytes;  
RX MEDLINE=93162043; PubMed=1286667;  
RA Rasmussen H.H., van Damme J., Puype M., Gesser B., Cellis J.E.,  
Vandekerckhove J.;  
RT "Microsequences of 145 proteins recorded in the two-dimensional gel  
protein database of normal human epidermal keratinocytes.";  
RL Electrophoresis 13:960-969(1992).  
[8]  
SEQUENCE OF 5-34.  
RX MEDLINE=93139333; PubMed=8423249;  
RA Miyasaki K.T., Bodeau A.L., Murthy A.R., Lehrer R.I.;  
RT "In vitro antimicrobial activity of the human neutrophil cytosolic S-  
100 protein complex, calprotectin, against Capnocytophaga sputigena.";

Mon Mar 8 10:18:08 2004

```

J. Dent. Res. 72:517-523(1993).
-!- FUNCTION: Expressed by macrophages in acutely inflamed tissues
and in chronic inflammations. Seem to be an inhibitor of protein
kinases. Also expressed in epithelial cells constitutively or
induced during dermatoses. May interact with components of the
intermediate filaments in monocytes and epithelial cells.
-!- MISCELLANEOUS: Has been shown to bind calcium.
-!- SIMILARITY: Belongs to the S-100 family.
-!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; X06233; CAA29579.1; -
EMBL; M21064; AAA36326.1; -
EMBL; M26311; AAA68480.1; -
EMBL; AF237581; AAF62536.1; -
EMBL; AF237582; AAF62537.1; -
EMBL; A12029; CAA01002.1; -
EMBL; A12032; CAA01004.1; -
EMBL; BC047681; AAH47681.1; -
PIR; B31848; B31848.
PDB; 1IRJ; 14-JAN-03.
SWISS-2DPAGE; P06702; HUMAN
Aarhus/Ghent-2DPAGE; 5007; IEF.
Aarhus/Ghent-2DPAGE; 6010; IEF.
Aarhus/Ghent-2DPAGE; 6017; IEF.
Aarhus/Ghent-2DPAGE; 7013; IEF.
PMAA-2DPAGE; P06702; -
Gnew; HGNC:10499; S100A9.
MIM; 123886; -
GO; GO:0005509; F:calcium ion binding; TAS.
GO; GO:0004871; F:signal transducer activity; TAS.
GO; GO:0007267; P:cell-cell signaling; TAS.
GO; GO:0006954; P:inflammatory response; TAS.
InterPro; IPR001751; CaBP_S100.
InterPro; IPR002048; EF-hand.
Pfam; PF00036; ehand; 1.
Pfam; PF01023; S_100; 1.
ProDom; PD003407; CaBP_S100; 1.
PROSITE; PS00018; EF_HAND; FALSE_NEG.
PROSITE; PS00303; S100_CABP; 1.
Calcium-binding; Macrophage; Phosphorylation; Polymorphism;
3D-structure. 23 36 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
CA_BIND 67 78 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
CA_BIND 113 113 PHOSPHORYLATION.
MOD_RES 20 20 H -> R.
VARIANT 20 20 /FTID-VAR_013008.
CONFLICT 6 6 S -> H (IN REF. 8).
CONFLICT 25 25 K -> F (IN REF. 8).
CONFLICT 28 28 H -> L (IN REF. 8).
CONFLICT 28 28 H -> L (IN REF. 8).
SEQUENCE 114 AA; 13242 MW; C3BE19729E14C078 CRC64;
Query Match 45.7%; Score 213.5; DB 1; Length 114;
Best Local Similarity 45.1%; Pred. No. 3.3e-13;
Matches 41; Conservative 25; Mismatches 24; Indels 1; Gaps 1;
1 TKLEHLEGIINIFHOYSVRLGHVDYTLIKRELKQLITKELPNTL-KNTKQGTIDKIFQN 59
6 SQLESIETIINTHOYSVKLGHPDTLNQGEFEKVLKQLNFKLNKQEKVIEHIMED 65
QY
DB
QY 60 LDANQEQVSFKFEVVLVTDVLTADNHIK 90
DB 66 LTNQDNQLSFEBEIMMARLTWASHEKME 96

```







```
DR GO; GO:0042035; P:regulation of cytokine biosynthesis; ISS.  
DR GO; GO:0048169; P:regulation of long-term neuronal synaptic p...; ISS.  
DR GO; GO:0006417; P:regulation of protein biosynthesis; ISS.  
DR InterPro; IPR001751; CAbP S100.  
DR InterPro; IPR002048; EF-hand.  
DR Pfam; PF00036; ehand; 1.  
DR Pfam; PF01023; S_100; 1.  
DR ProDom; PD003407; CAbP_S100; 1.  
DR PROSITE; PS00018; EF_HAND; 1.  
DR PROSITE; PS00303; SLF_CABP; 1  
KW Calcium-binding; Zinc; Metal-binding; 3D-structure.  
FT INIT_MET 0  
FT CA_BIND 18 31 EF-HAND 1 (LOW AFFINITY).  
FT CA_BIND 61 72 EF-HAND 2 (HIGH AFFINITY).  
FT HELIX 2 17  
FT TURN 18 19  
FT STRAND 27 27  
FT HELIX 29 39  
FT TURN 41 42  
FT HELIX 43 46  
FT HELIX 50 61  
FT TURN 62 62  
FT STRAND 68 68  
FT HELIX 70 83  
FT TURN 87 88  
SQ SEQUENCE 91 AA; 10613 MW; 2378AA8B8FF7134D CRC64;  
  
Query Match 35.1%; Score 164; DB 1; Length 91;  
Best Local Similarity 37.4%; Pred. No. 9.Se-09:  
Matches 34; Conservative 19; Mismatches 38; Indels 0; Gaps 0;  
  
QY 1 TKLEDHLEGLINFHQSVLGVLDYLKRELKLITKEPLTKNTKDQGTDIKIFQL 60  
Db :|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||::|||:  
1 SELERKANVALIDVHFQVGREGDKHLKKSELKELINNLSHPLEEIKEQVDVKWEIL 60  
  
QY 61 DANQDSQVSFKFPVLVTDLTAHNHIKE 91  
Db ::::::::::::|::|:  
61 DEDGGECDPFEPMAFMVMVTTACHEFFEHE 91  
  
RESULT 14  
S1OZ HUMAN STANDARD; PRt; 98 AA.  
ID _S1OZ_HUMAN STANDARD; PRt; 98 AA.  
AC Q8WXGB;  
DT 28-FEB-2003 (Rel. 41, Created)  
DD 28-FEB-2003 (Rel. 41, Last sequence update)  
DI 10-OCT-2003 (Rel. 42, Last annotation update)  
DE S-100Z protein.  
GN GN S100Z.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mamalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
OX NCBI_TaxId=9606;  
[1]  
SEQUENCE FROM N.A.; SEQUENCE OF 1-20, SUBUNIT, CALCIUM-BINDING,  
TISSUE SPECIFICITY, AND INTERACTION WITH S10OP.  
RP TISSUE-Prostate;  
RX MEDLINE=21614385; PubMed=11747429;  
R Gripenko A.V., Hopper J.E., Makhatadze G.I.;  
RT "Molecular characterization and tissue distribution of a novel member  
of the S100 family of EF-hand proteins.";  
RL Biochemistry 40:15538-15548(2001).  
[2]  
SEQUENCE FROM N.A.  
RP TISSUE=Testis;  
RC MEDLINE=22388257; PubMed=12477932;  
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buotow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.P., Jordan H., Moore T.J., Max S.I., Wang J., Hsieh P.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.
```



0b 2 SELEKALSNIIDVYHNISNIOGNHHALYKNDPKMVTTECPQFVQNI-----NIENLFREL 57

2y 61 DANQDEQVSEKFEVVLVTDVLITAHNDNIHKE 91

0b 58 DINSNDNAINFEFLAMVIKGVASHKOSHKE 88

Search completed: March 5, 2004, 09:47:10  
Job time : 15 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

3M protein - protein search, using sw model

Run on: March 5, 2004, 09:43:08 ; Search time 39 Seconds  
(without alignments) 736.209 Million cell u

Title: US-09-646-651C-1

perfect score: 467

```
sequence: 1 TKLEDHLEGIINIFHQYSVR.....EFVVLVTDVLITAHDNIIHKE 91
```

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters:

```
Minimum DB seq length: 0
```

000000000007:UJbUat fap vt mmmtyx.

```
Post-processing: Minimum Match 0%
```

Maximum Match 100%  
Listing first 45 summaries

Database :

```

1:  sp_archaea:*
2:  sp_bacteria:*
3:  sp_fungi:*
4:  sp_human:*
5:  sp_invertebrate:*
6:  sp_mammal:*
7:  sp_mhc:*
8:  sp_organelle:*
9:  sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vtebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaea:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result	No.	Score	Match	Query %	Length	DB	ID	Description
1	289	61.9	70	6	Q9TR16		Q9tr16 bos taurus	
2	193	41.3	100	13	Q7ZVA4		Q7zva4 brachydani	
3	178	38.1	99	13	Q8AVT2		Q8ayj2 squalus ac	
4	178	38.1	101	13	Q93395		Q93395 salvelinus	
5	167	35.8	92	11	Q925T3		Q925t3 cricetulus	
6	164	35.1	92	13	Q9PSF6		Q9psf6 ictalurus p	
7	163	34.9	94	11	Q9D3M4		Q9d3m4 mus muscul	
8	163	34.9	94	11	Q91V77		Q91v77 m 11 days e	
9	162	34.7	79	11	Q9JL08		Q9j108 mus muscul	
10	154	33.0	148	11	Q8BLX1		Q8blx1 mus muscul	
11	154	33.0	2496	11	Q8VHD8		Q8vhd8 mus muscul	
12	151	32.3	591	4	Q01720		Q01720 homo sapien	
13	151	32.3	687	4	Q9H4U2		Q9h4u2 homo sapien	
14	148	31.7	101	6	Q9TV56		Q9tv56 canis fami	
15	147	31.5	1218	4	Q05331		Q05331 homo sapien	
16	145	31.0	213	4	Q9H4U1		Q9h4u1 homo sapien	

17	140.5	30.1	83	11	Q9QVR5	Q9QVR5 rattus sp.
18	139.5	29.9	79	6	Q865V3	Q865V3 equus caball
19	139	29.8	103	6	Q862H7	Q862H7 bos taurus
20	132	28.3	501	11	Q8CIU0	Q8CIU0 rattus norv
21	129	27.6	357	11	Q8BHC3	Q8BHC3 mus musculus
22	129	27.6	554	11	Q8K5S2	Q8K5S2 mus musculus
23	128	27.4	89	11	Q9R2B7	Q9R2B7 rattus norv
24	127.5	27.3	217	13	Q9YH57	Q9YH57 rana catesbe
25	126	27.0	80	11	Q91XG5	Q91XG5 mus musculus
26	124	26.6	495	4	Q9UBG3	Q9UBG3 homo sapien
27	124	26.6	495	4	Q8N6L3	Q8N6L3 homo sapien
28	112.5	24.1	73	6	Q28714	Q28714 cryotolagus
29	112	24.0	77	6	Q862V4	Q862V4 bos taurus
30	105	22.5	55	6	Q9TFSB1	Q9TFSB1 bos taurus
31	101.5	21.7	124	11	Q9D708	Q9D708 mus musculus
32	90.5	19.4	170	13	Q7T0G3	Q7T0G3 xenopus lae
33	88	18.8	638	11	Q9D3P1	Q9D3P1 mus musculus
34	87.5	18.7	170	5	Q9NKW7	Q9NKW7 parinopecte
35	84.5	18.1	535	10	Q38873	Q38873 arabidopsis
36	82	17.6	110	4	Q9Y544	Q9Y544 homo sapien
37	82	17.6	169	5	Q9NFN1	Q9NFN1 schistosoma
38	81.5	17.5	536	10	Q65644	Q65644 arabidopsis
39	81	17.3	160	5	Q9BUE9	Q9BUE9 plasmodium
40	81	17.3	170	5	Q9SP81	Q9SP81 bombyx mori
41	80.5	17.2	248	10	Q65587	Q65587 arabidopsis
42	79.5	17.0	512	10	Q9SE34	Q9SE34 oryza sativ
43	78.5	16.8	115	11	Q99LQ9	Q99LQ9 mus musculus
44	78.5	16.8	171	5	Q20804	Q20804 caenorhabdi
45	78.5	16.8	523	10	Q9FKW4	Q9FKW4 arabidopsis

## ALIGNMENTS

## RESULT 1

Q9TR16	Q9TR16	PRELIMINARY;	PRT;	70 AA.
ID	Q9TR16			
AC	Q9TR16;			
DT	01-MAY-2000 (TrEMBLrel. 13, Created)			
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)			
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)			
DE	CORNEA-associated antigen, CO-AG-CALGRANULIN C homolog.			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Bovinae; Bos.			
OX	NCBI TaxID=9913;			
OX	[1]			
RP	SEQUENCE.			
RP				
EX	MEDLINE=96181454; PubMed=8603881;			
RA	Liu S.H., Gottsch J.D.;			
RA	"Amino acid sequence of an immunogenic corneal stromal protein.";			
RL	Invest. Ophthalmol. Vis. Sci. 37:944-948(1996).			
CC	!- SIMILARITY: BELONGS TO THE S-100 FAMILY.			

## Description

```

DR  GO: 0005509; F:calcium ion binding; IEA.
DR  InterPro: IPR001751; CaBP S100.
DR  InterPro: IPR002048; EF-hand.
DR  Pfam: PF01023; S_100; 1.
DR  ProDom: PD003407; CaBP S100; 1.
SQ SEQUENCE 70 AA; 8134 NW; 7D52BEA97A4D53A5 CRC64;

Query Match      61.9%; Score 289; DB 6; Length 70;
Best Local Similarity 81.4%; Pred. No. 3.3e-20;
Matches 57; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

```

[illegible]

Qy 61 DANODEOVSE 70

Db 61 DADKKGAVVF 70

DR PROSITE; PS00018; EF\_HAND; 1.  
DR PROSITE; PS00303; S100\_CABP; 1.  
SQ SEQUENCE 99 AA; 11050 MW; BA62D8190A4A3693 CRC64;

Query Match 38.1%; Score 178; DB 13; Length 99;  
Best Local Similarity 40.7%; Pred. No. 1.7e-09;  
Matches 35; Conservative 20; Mismatches 31; Indels 0; Gaps 0;

QY 1 TKLEHLEGIINIFHQYSVRLGHVDTLTKRELKQLITKELPNTLKNTKDQGTIDKIFQNL 60  
Db 2 TELESAMAGILGVFRKSGKGGDKYSLSNNEMVDLLKALPNFLKSKQDKAAVDTIMKDL 61  
QY 61 DANQDEQVSFKFVVLVTDVLIATAD 86  
Db 62 DRNKGELDFQFVVLIAALAAACND 87

RESULT 4  
O93395 PRELIMINARY; PRT; 101 AA.  
AC O93395;  
DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE S100-like calcium binding protein.  
GN S100.  
OS Salvelinus fontinalis (Brook trout) (Brook char).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.  
OX NCBI\_TaxID=8038;  
RN [1] SEQUENCE FROM N.A.

RP TISSUE=Ovary;  
RC MEDLINE=20534789; PubMed=11080585;  
RA Bobe J., Goetz F.W.;  
RT "A S100 homologue mRNA isolated by differential display PCR is down-regulated in the brook trout (Salvelinus fontinalis) post-ovulatory ovary."  
RT ovary."  
RL Gene 257:187-194(2000).  
CC -!- SIMILARITY: BELONGS TO THE S-100 FAMILY.  
DR EMBL; AF077613; AAC28367.1; -.  
DR HSP; P02633; 4ICB.  
DR GO; GO:0005509; F:calcium ion binding; IEA.  
DR InterPro; IPR001751; CaBP S100.  
DR InterPro; IPR002048; EF-hand.  
DR Pfam; PF00036; ehand; 1.  
DR Pfam; PF01023; S100; 1.  
DR ProDom; PD003407; CaBP S100; 1.  
DR ProDom; PD000012; EF-hand; 1.  
DR PROSITE; PS00018; EF\_HAND; 1.  
DR PROSITE; PS00303; S100\_CABP; 1.  
DR Hypothetical protein.  
SQ SEQUENCE 100 AA; 11157 MW; DAB81814E54CB8C CRC64;

Query Match 41.3%; Score 193; DB 13; Length 100;  
Best Local Similarity 46.1%; Pred. No. 6.3e-11;  
Matches 41; Conservative 17; Mismatches 31; Indels 0; Gaps 0;

QY 3 LEDHLEGIINIFHQYSVRLGHVDTLTKRELKQLITKELPNTLKNTKDQGTIDKIFQNLDA 62  
Db 5 LERAMETLITVFRHYSGAGNSSTLSRELKQLMEKELASFLKSKQDPAAVDKIMKDLDA 64  
QY 63 NQDEQVSFKFVVLVTDVLIATADNIHKE 91  
Db 65 NGDGEVNFEEFVSLVGLSLACQLYQKQ 93

RESULT 3  
Q8AYJ2 PRELIMINARY; PRT; 99 AA.  
AC Q8AYJ2;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE S-100 calcium-binding protein Al.  
OS Squalus acanthias (Spiny dogfish).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;  
OC Elasmobranchii; Squalia; Squaloidei; Squalidae; Squalus.  
OX NCBI\_TaxID=7797;  
RN [1] SEQUENCE FROM N.A.

RP Wang C., Callard G.V.;  
RC "Molecular cloning and stage dependence of an S-10 cDNA from the shark testis."  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF421551; AAN63527.1; -.  
DR GO; GO:0005509; F:calcium ion binding; IEA.  
DR InterPro; IPR001751; CaBP S100.  
DR InterPro; IPR002048; EF-hand.  
DR Pfam; PF00036; ehand; 1.  
DR Pfam; PF01023; S100; 1.  
DR ProDom; PD003407; CaBP S100; 1.  
DR ProDom; PD000012; EF-hand; 1.  
DR SMART; SM00054; EFH; 1.

```
01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
SI00B.
Cricetus griseus (Chinese hamster).
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
Cricetulus.
NCBI_TaxID=10029;
[1]
IN SEQUENCE FROM N.A.
Kuge O., Yamakawa Y., Nishijima M.;
"Enhancement of transport-dependent decarboxylation of
phosphatidylserine by SI00B protein in permeabilized Chinese hamster
ovary cells.";
J. Biol. Chem. 0:0-0(2001).
EMBL; AB056121; BAB43945.1; --
-!- SIMILARITY: BELONGS TO THE S-100 FAMILY.
GO; GO:0005737; C:cytoplasm; ISS.
GO; GO:0005576; C:extracellular; ISS.
GO; GO:0005509; F:calcium ion binding; ISS.
GO; GO:0019210; F:kinase inhibitor activity; ISS.
GO; GO:0042803; F:protein homodimerization activity; ISS.
GO; GO:0048155; F:S100 alpha binding; ISS.
GO; GO:0048154; F:S100 beta binding; ISS.
GO; GO:0048156; F:tau protein binding; ISS.
GO; GO:0008270; F:zinc ion binding; ISS.
GO; GO:0048143; F:astrocyte activation; ISS.
GO; GO:0007409; P:axonogenesis; ISS.
GO; GO:0006874; P:calcium ion homeostasis; ISS.
GO; GO:0006112; P:energy reserve metabolism; ISS.
GO; GO:0048151; P:hyperphosphorylation; ISS.
GO; GO:0006917; P:induction of apoptosis; ISS.
GO; GO:0007511; P:learning and/or memory; ISS.
GO; GO:0045917; P:positive regulation of complement activation; ISS.
GO; GO:0042035; P:regulation of cytokine biosynthesis; ISS.
GO; GO:0048169; P:regulation of long-term neuronal synaptic p. .; ISS.
GO; GO:0006417; P:regulation of protein biosynthesis; ISS.
InterPro; IPR001751; CaBP_S100.
InterPro; IPR002048; EF-hand.
PFam; PF00036; efhand; 1.
PFam; PF01023; S_100; 1.
ProDom; PD003407; CaBP_S100; 1.
PROSITE; PS00018; EF_HAND; 1.
PROSITE; PS00303; S100_CaBP; 1.
SEQUENCE 92 AA; 10749 MW; AF50107EC2BEDF6B CRC64;

Query Match 35.8%; Score 167; DB 11; Length 92;
Best Local Similarity 38.5%; Pred. No. 1.7e-08;
Matches 35; Conservative 19; Mismatches 37; Indels 0; Gaps 0;

1 TKLEHLEGIINPHQYSVRLGHVDTLTKELKQLITKELPNTLKNKQGTIDKIFONL 60
::: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2 SELEKAMVALIDIFHQYSGEGDKHKKSELKELINNELSHLEETKEQEVVDKVMETL 61
::: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

61 DANQEQVSFKPVVLTVDLTAHNDNIHKE 91
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
62 DEDGEGCEDFOEFMAFVSMVTTCHEFFQE 92
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 6
Q9PSF6 PRELIMINARY; PRT; 92 AA.
01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
Ictaluridae; Ictalurus.
NCBI_TaxID=7998;
[1]
IN SEQUENCE FROM N.A.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
Ictaluridae; Ictalurus.
NCBI_TaxID=7998;
[1]
IN SEQUENCE FROM N.A.
```

```
RX MEDLINE=94376615; PubMed=8090068;
RA Bettini E., Porta A.R., Dahmen N., Wang H., Margolis F.L.;
RT "Expressed sequence tags (EST) identify genes preferentially expressed
in catfish chemosensory tissues.";
RL Brain Res. Mol. Brain Res. 23:285-291(1994).
CC -!- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR HSP; P30801; 1A03.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR InterPro; IPR001751; CaBP_S100.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; efhand; 1.
DR Pfam; PF01023; S_100; 1.
DR ProDom; PD003407; CaBP_S100; 1.
DR ProDom; PD000012; EF-hand; 1.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CaBP; 1.
SQ SEQUENCE 92 AA; 10022 MW; E04875D0C9921C50 CRC64;

Query Match 35.1%; Score 164; DB 13; Length 92;
Best Local Similarity 41.5%; Pred. No. 3.2e-08;
Matches 34; Conservative 15; Mismatches 33; Indels 0; Gaps 0;

10 IINIFHOYSVRLGHVDTLTKELKQLITKELPNTLKNKQGTIDKIFONLDAQBQVS 69
::: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
11 LISTFHKYSGEEDKCTLTGELKDLTKELGAFGNCSDQATLTKIFOLDTNADGVVD 70
::: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

70 FK3FVVLVDTVLTAHNDNIHKE 91
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
71 FOEYATWACTTMCNKSLSKK 92
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 7
Q9D3M4 PRELIMINARY; PRT; 94 AA.
01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE S100 calcium binding protein A1.
GN S100A1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RV [1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Harai A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yananaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nardone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -!- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR EMBL; AK017279; BAB30670.1; --
DR HSP; P04631; 1B4C.
DR MGD; MGI:1338917; S100a1.
DR GO; GO:0005509; F:calcium ion binding; IEA.
```











GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

M nucleic - nucleic search, using sw model  
Run on: March 8, 2004, 03:14:22 ; Search time 2062 Seconds  
(without alignments)  
2249.129 Million cell updates/sec  
Title: US-09-646-651c-2  
Perfect score: 107  
Sequence: 1 ggaagaaunnnunnaugnn.....nnnnnnnnnnnnuagcag 107

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 3470272 seqs, 21671516995 residues 6940544  
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*  
29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pln.\*  
35: em.htg.tod.\*  
36: em.htg.mam.\*  
37: em.htg.vrt.\*  
38: em.sy.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	31.4	29.3	298	6	A94582	A94582 Sequence 3
2	30	28.0	162986	9	CNS051BQ	AL163953 Human chr
3	30	28.0	167398	9	CNS07BET	AL445886 Human chr
4	29.8	27.9	1311	3	BT001281	BT001281 Drosophila
5	29.8	27.9	133636	9	AC117436	AC117436 Homo sapi
6	29.6	27.7	86050	9	AL359647	AL359647 Human DNA
7	29.4	27.5	17492	1	U67539	U67539 Methanococc
8	29.4	27.5	110000	6	AR271569_05	Continuation (6 of
9	29.2	27.3	60471	2	AC015173	AC015173 Drosophila
10	29.2	27.3	164038	3	AC009846	AC009846 Drosophila
11	29.2	27.3	234378	3	AE003740	AE003740 Drosophila
12	29	27.1	191942	9	AC079951	AC079951 Homo sapi
13	28.6	26.7	25233	3	CEC01B9	Z93373 Caenorhabdi
14	28.6	26.7	113880	3	PFMAL3P4	AL008970 Plasmodiu
15	28.6	26.7	227182	2	AC099283	AC099283 Rattus no
16	28.6	26.7	252129	2	AC133975	AC133975 Rattus no
17	28.4	26.5	1818	1	AF267206	AF267206 Candidatu
18	28.4	26.5	112537	9	AC112251	AC112251 Homo sapi
19	28.4	26.5	167560	7	AY303349	AY303349 Enterobac
20	28.4	26.5	177897	2	AC021473	AC021473 Homo sapi
21	28.4	26.5	178204	9	AC097467	AC097467 Homo sapi
22	28.2	26.4	49843	3	AF288092	AF288092 Naegleria
23	28	26.2	146500	9	AC004453	AC004453 Homo sapi
24	28	26.2	216941	2	AC146064	AC146064 Pan trogl
25	28	26.2	263150	2	AC105564	AC105564 Rattus no
26	27.8	26.0	150641	3	CEV41E3	Z95559 Caenorhabdi
27	27.8	26.0	178607	9	AC007058	AC007058 Homo sapi
28	27.8	26.0	181542	2	AC121076	AC121076 Canis fam
29	27.6	25.8	102704	5	AL672216	AL672216 Zebrafish
30	27.6	25.8	120039	9	AL357124	AL357124 Human DNA
31	27.6	25.8	132254	3	AC116330	AC116330 Dictyoste
32	27.6	25.8	147341	8	AC122143	AC122143 Oryza sat
33	27.6	25.8	160404	9	AL159152	AL159152 Human DNA
34	27.6	25.8	166315	9	AC078940	AC078940 Homo sapi
35	27.4	25.6	524	6	AX541033	AX541033 Sequence
36	27.4	25.6	769	6	AX541172	AX541172 Sequence
37	27.4	25.6	789	6	AX541169	AX541169 Sequence
38	27.4	25.6	835	6	AX541247	AX541247 Sequence
39	27.4	25.6	110000	2	PFMAL13_11	Continuation (12 o
40	27.4	25.6	138027	9	AL590227	AL590227 Human DNA
41	27.4	25.6	152313	9	AL591594	AL591594 Human DNA
42	27.4	25.6	173342	9	AC023796	AC023796 Homo sapi
43	27.4	25.6	183602	2	AC024025	AC024025 Homo sapi
44	27.4	25.6	195583	9	AC092867	AC092867 Homo sapi
45	27.4	25.6	283919	2	AC095169	AC095169 Rattus no

ALIGNMENTS

RESULT 1  
A94582  
LOCUS A94582 298 bp DNA linear PAT 26-JAN-2000  
DEFINITION Sequence 3 from Patent WO9947561.  
ACCESSION A94582  
VERSION A94582.1 GI:6778899  
KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1 (bases 1 to 298)  
AUTHORS Koch-Pelster, B. and Kuhn, E.  
TITLE RIBONUCLEOTIDE POLYPEPTIDES CONTAINING METAL  
JOURNAL Patent: WO 9947561-A 3 23-SEP-1999;  
Koch Pelster Brigitte (DE); Kuhn Eckehard (DE)

	FEATURES	Location/Qualifiers
	source	1..162986
		/organism="Homo sapiens"
		/mol_type="genomic DNA"
		/db_xref="taxon:9606"
		/chromosome="14"
		/clone="R-533L7"
		/clone_lib="RPCI-11"
	SFS	133879..134005
		/note="matching EMBL:N74379
		RHdb:RH91341
		dbSTS:STS64476
		Identified using the e-PCR software (G. Schuler)"
	SFS	151705..151879
		/note="matching EMBL:G03450
		RHdb:RH53483
		RHdb:RH3638
		dbSTS:STS817
		Identified using the e-PCR software (G. Schuler)"
	ORIGIN	
		Query Match 28.0%; Score 30; DB 9; Length 162986;
		Best Local Similarity 22.1%; Pred. No. 1.3e+02;
		Matches 23; Conservative 13; Mismatches 68; Indels 0; Gaps 0;
	QY	3 AAAAUNNNNNUAAUGNNNNNCUNUUUUNNNNNNAAAAUAUAAACAUNNNNNCUU 62
	Db	159690 AAAAAATCATATTGTATTAATCTTTTTCCTGGTGACAGGCTAGAGAAAAGGATGCTT 159631
	QY	63 NAGNNNNNNNNNNAGAAUNNNNNNNNNNNNNNNNNNNNUUAGCA 106
	Db	159630 GGGTGGCAATTGAATAAGAAATCCACTTCAAATACATATTACCA 159587
	RESULT 3	
	CNS07BET/c	
	LOCUS	CNS07BET 167398 bp DNA linear PRI 29-APR-2001
	DEFINITION	Human chromosome 14 DNA sequence BAC R-369B6 of library RPCI-11 from chromosome 14 of Homo sapiens (Human), complete sequence.
	ACCESSION	AL445886
	VERSION	AL445886.3 GI:13897491
	KEYWORDS	HTG.
	SOURCE	Homo sapiens (human)
	ORGANISM	Homo sapiens
		Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
	REFERENCE	1 (bases 1 to 167398)
	AUTHORS	Heilig,R., Petit,J.L., Vico,V., Dasilva,C., Robert,C., Wincker,P., Brottier,F., Cattolico,L., Barbe,V., Pelletier,E., Artiguenave,F., Levy,M., Eckenberg,R., Bruls,T., deBerardinis,V., Cruaud,C., Gyapay,G., Saurin,W. and Weissenbach,J.
		Gyapay,G., Saurin,W. and Weissenbach,J.
		Sequencing of the human chromosome 14
	TITLE	Unpublished
	JOURNAL	2 (bases 1 to 167398)
	REFERENCE	Genoscope.
	AUTHORS	Direct Submission
	TITLE	Submitted (02-MAY-2001) Genoscope - Centre National de Sequencage :
	JOURNAL	BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
		- Web : www.genoscope.cns.fr)
	COMMENT	On Apr 30, 2001 this sequence version replaced gi:11181947. ----- Genome Center Center code: GS Web site: http://www.genoscope.cns.fr/ Contact: SeqRef@genoscope.cns.fr
		The following BAC sequence is oriented from the T7 to the SP6 end.
		----- Summary Statistics
		Assembly program: Phrap; version 2.0
		Quality coverage: 8.81x in Q20 bases; sum-of-contigs
		-----
		Overall quality chart :
		Range : bases
		0 : 9
		10 - 19 : 2
		20 - 29 : 2
		30 - 39 : 2
		40 - 49 : 251
		50 - 59 : 1196
		60 - 69 : 2599
		70 - 79 : 9911
		80 - 89 : 40597
		90 - 99 : 108430
		-----
		Percentage of bases with a quality value >= 40 : 99 %
	FEATURES	Location/Qualifiers
	source	1..298
		/organism="unidentified"
		/mol_type="unassigned DNA"
		/db_xref="taxon:32644"
	ORIGIN	
		Query Match 29.3%; Score 31.4; DB 6; Length 298;
		Best Local Similarity 30.3%; Pred. No. 3.7e+02;
		Matches 23; Conservative 9; Mismatches 44; Indels 0; Gaps 0;
	QY	32 UUNNNNNAAAAAUAAACAUNNNNNCUNAGNNNNNNNNNNNNNNNNAGAAUUNNNNN 91
	Db	89 TTGTGCATATAAATAAACATGACTTCTTTAGACACTCTTCATTAGAAATAAATAA 148
	QY	92 NNNNNNNNUAGCAG 107
	Db	149 AATAAACTATTAGCAG 164
	RESULT 2	
	CNS05TBQ/c	
	LOCUS	CNS05TBQ 162986 bp DNA linear PRI 03-MAY-2001
	DEFINITION	Human chromosome 14 DNA sequence BAC R-533L7 of library RPCI-11 from chromosome 14 of Homo sapiens (Human), complete sequence.
	ACCESSION	AL163953
	VERSION	AL163953.3 GI:12001746
	KEYWORDS	HTG.
	SOURCE	Homo sapiens (human)
	ORGANISM	Homo sapiens
		Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
	REFERENCE	1 (bases 1 to 162986)
	AUTHORS	Heilig,R., Petit,J.L., Vico,V., Dasilva,C., Robert,C., Wincker,P., Brottier,F., Cattolico,L., Barbe,V., Pelletier,E., Artiguenave,F., Levy,M., Eckenberg,R., Bruls,T., deBerardinis,V., Cruaud,C., Gyapay,G., Saurin,W. and Weissenbach,J.
		Gyapay,G., Saurin,W. and Weissenbach,J.
		Sequencing of the human chromosome 14
	TITLE	Unpublished
	JOURNAL	2 (bases 1 to 162986)
	REFERENCE	Genoscope.
	AUTHORS	Direct Submission
	TITLE	Submitted (02-MAY-2001) Genoscope - Centre National de Sequencage :
	JOURNAL	BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
		- Web : www.genoscope.cns.fr)
	COMMENT	On Jan 1, 2001 this sequence version replaced gi:7768128. ----- Genome Center Center code: GS Web site: http://www.genoscope.cns.fr/ Contact: SeqRef@genoscope.cns.fr
		The following BAC sequence is oriented from the T7 to the SP6 end.
		----- Summary Statistics
		Assembly program: Phrap; version 2.0
		Quality coverage: 8.81x in Q20 bases; sum-of-contigs
		-----
		Overall quality chart :
		Range : bases
		0 : 9
		10 - 19 : 2
		20 - 29 : 2
		30 - 39 : 2
		40 - 49 : 251
		50 - 59 : 1196
		60 - 69 : 2599
		70 - 79 : 9911
		80 - 89 : 40597
		90 - 99 : 108430
		-----
		Percentage of bases with a quality value >= 40 : 99 %



Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

1 (bases 1 to 133636)  
REFERENCE  
AUTHORS  
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C.,  
Alsbrooks, S.L., Anwarungrasame, H.C., Are, J.R., Ayale, M., Banks, T.,  
Barbieri, J., Benton, J., Binage, K., Blankenburg, K., Bonnin, D.,  
Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,  
Cahoon, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,  
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,  
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,  
Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,  
Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,  
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,  
Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,  
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,  
Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,  
Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,  
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., He, X.,  
Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C.,  
Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J.,  
Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S.,  
Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J.,  
Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, B.C.,  
Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W.,  
Loulsegh, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R.,  
Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A.,  
Martinez, E., Massey, E., Mawhinney, E., McLeod, M.P., Meador, M.,  
Mei, G., Metzger, M., Miner, Z., Miner, Z., Mitchell, T., Mohabbat, K.,  
Moore, S., Morgan, M., Moorish, T., Morris, S., Moser, M., Neal, D.,  
Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N.,  
Nickerson, E., Nwokenkwo, S., Oguh, M., Okwuonu, G., Oragunye, N.,  
Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L.,  
Pickens, R., Primus, E., Pu, D.L., Quiles, M., Ren, Y., Rives, M.,  
Rojas, A., Rotobokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S.,  
Scott, G., Shen, H., Shooshitari, N., Sisson, I., Sodergren, E.,  
Sonaite, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A.,  
Tabors, P., Tanerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C.,  
Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L.,  
Vera, V., Villalón, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S.,  
Warren, R., Washington, C., Watlington, S., Williams, G.,  
Williamson, A., Wleczek, R., Wooden, S., Worley, K., Wu, C., Wu, Y.,  
Wu, Y.F., Zhou, J., Zorrilla, S., Naylor, S.L., Weinstein, G. and  
Gibbs, R.

Direct Submission  
Unpublished  
2 (bases 1 to 133636)  
Worley, K.C.

Direct Submission  
Submitted (10-APR-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 133636)  
Worley, K.C.

Direct Submission  
Submitted (29-JUN-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

4 (bases 1 to 133636)  
Worley, K.C.

Direct Submission  
Submitted (30-SEP-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

5 (bases 1 to 133636)  
Worley, K.C.

Direct Submission  
Submitted (01-OCT-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

On Sep 30, 2002 this sequence version replaced gi:21553164.  
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email  
[gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

#### ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:  
<http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.ht>

ml. Location/Qualifiers

1. 133636  
/organism="Homo sapiens"  
/db\_type="genomic DNA"  
/mol\_xref="taxon:9606"  
/chromosome="3"  
/clone="RP11-69K16"  
7. 197  
/standard\_name="RH122785"  
complement(513..810)  
/rpt\_family="AluSx"  
833..1131  
/rpt\_family="AluJb"  
1423..1453  
/rpt\_family="AT-rich"  
1599..1747  
/rpt\_family="MER46C"  
1794..2232  
/rpt\_family="MER9"  
2233..2253  
/rpt\_family="(TG)n"  
2254..2310  
/rpt\_family="MER9"  
8037..8312  
/rpt\_family="HERVK\_9"  
8313..8845  
/rpt\_family="MER9"  
complement(8886..9207)  
/rpt\_family="L1M4C"  
9225..9396  
/rpt\_family="(TA)n"  
complement(9405..9825)  
/rpt\_family="L1M4C"  
complement(10094..10457)  
/rpt\_family="L1ME"  
complement(10456..10578)  
/rpt\_family="L1MEC"  
complement(10579..10881)  
/rpt\_family="AluJb"



```
complement(10882..11975)
/rpt_family="L1MEC"
11721..11845
/standard_name="SHGC-77614"
12033..12114
/rpt_family="MER46C"
12758..13289
/rpt_family="L2"
13532..13569
/rpt_family="AT_rich"
15616..15945
/rpt_family="L1PB3"
15946..15984
/rpt_family="(TAAA)n"
15985..17070
/rpt_family="L1PB3"
18445..19144
/rpt_family="L1PB3"
19145..19166
/rpt_family="AT_rich"
19232..19417
/rpt_family="L1M4"
19697..20414
/rpt_family="MER21B"
20465..20575
/rpt_family="L2"
20697..21004
/rpt_family="AluY"
22243..22552
/rpt_family="AluSq"
22553..22717
/rpt_family="AluSg/x"
23062..23121
/rpt_family="(TA)n"
23205..23829
/rpt_family="L2"
23799..24083
/rpt_family="L2"
24096..24150
/rpt_family="L2"
complement(24162..24377)
/rpt_family="MIR"
24378..24447
/rpt_family="L2"
24456..24795

Query Match      27.9%; Score 29.8; DB 9; Length 133636;
Best Local Similarity 23.4%; Pred. No. 1.6e+02;
Matches 25; Conservative 12; Mismatches 70; Indels 0; Gaps 0;

yy          1 GGAAAAUNNNNNUAUGNNNNNNNCUNNNUUUNNNNNNAANAAACUAUNNNNC 60
db          98804 GGAACAATCTATTATGTGACTTTTTCACATGCCCAAGAAATTTGAAGCATGGTCTC 98745

yy          61 UUNAGNNNNNNNNNNNNAGAAAAUNNNNNNNNNNNNNNNNNNNUUAAGCAG 107
db          98744 TGAATAAGCTTAAGACAGAAGAAATGGAAGGAATGGCCCTTTTGAG 98698

RESULT 6
AL359647
LOCUS              AL359647             86050 bp    DNA        linear   PRI 05-APR-2001
DEFINITION         Human DNA sequence from clone RP11-431P10 on chromosome 13,
                    complete sequence.
ACCESSION           AL359647
VERSION             AL359647.9  GI:13560015
KEYWORDS            HTG.
SOURCE              Homo sapiens (human)
ORGANISM            Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE            1 (bases 1 to 86050)
AUTHORS             Phillimore,B.
```

TITLE	JOURNAL	COMMENT
Direct Submission	Submitted (05-APR-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk	requests: clonerequest@sanger.ac.uk
	On Apr 6, 2001 this sequence version replaced gi:13273767.	
	During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.	
	This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; SW: SWISSPROT; Tr: TREML; Wp: WORMPEP; Information on the WORMPEP database can be found at <a href="http://www.sanger.ac.uk/Projects/C_elegans/wormpep">http://www.sanger.ac.uk/Projects/C_elegans/wormpep</a>	
	This sequence was generated from part of bacterial clone contigs of human chromosome 13, constructed by the Sanger Centre Chromosome 13 Mapping Group. Further information can be found at <a href="http://www.sanger.ac.uk/HGP/Chr13">http://www.sanger.ac.uk/HGP/Chr13</a>	
	RP11-431P10 is from the library RPC1-11.2 constructed by the group of Pieter de Jong. For further details see <a href="http://www.chori.org/bacpac/home.htm">http://www.chori.org/bacpac/home.htm</a>	
	VECTOR: pBACe3.6	
	IMPORTANT: This sequence is not the entire insert of clone RP11-431P10. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.	
	The true right end of clone RP11-431P10 is at 86050 in this sequence. The true left end of clone RP11-569C12 is at 48980 in this sequence. The true right end of clone RP11-600P1 is at 100 in this sequence.	

FEATURES	source	Location/Qualifiers
		1. .86050
		/organism="Homo sapiens"
		/mol_type="genomic DNA"
		/db_xref="taxon:9606"
		/chromosome="13"
		/clone="RP11-431P10"
		/clone_lib="RPC1-11.2"
misc_feature		110..1591
		/note="match: GSS: Em:AQ621405"
repeat_region		1507..1694
		/note="MERSA repeat: matches 1."
repeat_region		1836..2969
		/note="M1M4 repeat: matches 258"
repeat_region		2996..3141
		/note="M1M6C repeat: matches 14"
repeat_region		3270..4102
		/note="M1M6C repeat: matches 16"
repeat_region		4530..6107
		/note="M1PA15 repeat: matches 4"
repeat_region		7573..7622
		/note="25 copies 2 mer gt 78%"
repeat_region		8343..8653
		/note="AluXs repeat: matches 1."
repeat_region		9055..9356
		/note="M1M1E repeat: matches 25"
misc_feature		complement(9811..10118)
		/note="match: GSS: Em:AQ092782"
misc_feature		complement(10087..10508)
		/note="match: GSS: Em:AQ215557"
misc_feature		10231..10707
		/note="match: GSS: Em:AQ728730"
misc_feature		10260..10779
		/note="match: GSS: Em:AQ804059"
repeat_region		11538..11837



Fuhrmann, J.L., Nguyen, D.T., Utterback, T., Kelley, J.M., Peterson, J.D., Sadow, P.W., Hanna, M.C., Cotton, M.D., Hurst, M.A., Roberts, K.M., Kaine, B.B., Borodovsky, M., Klenk, H.P., Fraser, C.M., Smith, H.O., Woese, C.R., and Venter, J.C.  
Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*  
Science 273 (5278), 1058-1073 (1996)  
96337999  
8688087  
2 (bases 1 to 17492)  
Bult, C.J., White, O., Olsen, G.J., Zhou, L., Fleischmann, R.D., Sutton, G.G., Blake, J.A., FitzGerald, L.M., Clayton, R.A., Gocayne, J.D., Kerlavage, A.R., Dougherty, B.A., Tomb, J.-F., Adams, M.D., Reich, C.I., Overbeek, R., Kirkness, E.F., Weinstock, K.G., Merrick, J.M., Glodek, A., Scott, J.D., Geoghegan, N.S., Weidman, J.F., Fuhrmann, J.L., Nguyen, D.T., Utterback, T., Kelley, J.M., Peterson, J.D., Sadow, P.W., Hanna, M.C., Cotton, M.D., Hurst, M.A., Roberts, K.M., Kaine, B.B., Borodovsky, M., Klenk, H.P., Fraser, C.M., Smith, H.O., Woese, C.R. and Venter, J.C.  
Direct Submission  
Submitted (27-AUG-1996) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA  
On Oct 3, 1996 this sequence version replaced gi:15633994.  
Location/Qualifiers  
1. 17492  
/organism="Methanocaldococcus jannaschii"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:2190"  
complement(87..515)  
/gene="MJ0952"  
complement(87..515)  
/gene="MJ0952"  
/notes="similar to SP:Q04493 PID:575709 percent identity: 26.72; identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAB98957.1"  
/db\_xref="GI:1592316"  
TRANSLATION="MENARDLRKAMALEIYNQQLQIQSEITSIRALKSEIMNSIK TISNRKADETLIPVPGVFVKAKIVDDKALIGVKSIIYVEKSFNEVIEDLKKSVEDL DRAEKGMKKAELAKAITALRKEIQEIAQAQKQKQ"  
696..998  
/gene="MJ0953"  
696..998  
/gene="MJ0953"  
/notes="similar to GB:AE000666 percent identity: 52.00; identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAB98958.1"  
/db\_xref="GI:1499790"  
TRANSLATION="MRLLGVIGYLAVLTKAICESWVDVVKRSINGEIHQPVEIESII NNPTGLVLSWSITATPGTLVIDLIPEERKLKVAISPRREDIIVPEPEYIKKID"  
complement(1007..2329)  
/gene="MJ0954"  
complement(1007..2329)  
/gene="MJ0954"  
/notes="similar to GB:M59200 PID:150044 percent identity: 30.13; identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="S-layer structural protein"  
/protein\_id="AAB98959.1"  
/db\_xref="GI:1591621"  
TRANSLATION="MLRLQWMEGLIVKRTLILLVLSVALPIEPIIYNKSTVD YQNAKLMDFYSREININDNVTIINDIMYIPSDIELIKQKDKNLIIKFRDGN KVKYKDIETCYLKGKEISILENKYIVVEDITSYVILKEKDGKVELTNESEFYDG KVVVKLSDSLNTIIVDIYKNEKLSPLTKGKIYMKGGTGLGMYENCTRIQGY RETFRVYSTJKEGEDYPLDKFKVEISDPDKLEYNKIDSLGNEIYLYNTIPE KCYKDYFLVKIRKERTVDKVAIYGDGIYAVKVNNTVHFYAGKELKHEKIYLG SVDVYSNPLNVNKKIILIGGPKVNKIVKLEEDKGLKVNISTNYPGNNRGIILKIN

gene  
CDS  
PYNDNNIILAGSDRWGTAKAAILVFLTKYDNEDDTLMVEMDKGEIKIK"  
2378..3499  
/gene="MJ0955"  
2378..3499  
/gene="MJ0955"  
/note="similar to SP:P06986 GB:U02071 GB:X03416 PID:41695  
PID:41710 percent identity: 29.00; identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="histidinol-phosphate aminotransferase (hisC)"  
/protein\_id="AAB98960.1"  
/db\_xref="GI:1591622"  
TRANSLATION="MGDMWMIENKVDVVVKLKPYPVPGSKSEELIARAYGIKPEDIIKLG SNENPSPKKEKILDEIDIKHQYPEPVPNIMKBSKFLNVDENIIVGGGDAE IIDTFEFDVDDGEVLIPIPTFYRSATIHNAIKYAKYDKEDFKLNVSVLNN IITDKVIFLCTPNPTGNIENRDERVINETDALVIDHAYIEYAKKEYDTQKAP EYDNVILRTFSKVGAGMEVGVANKIIDYMMRVKPIPSLTLSOVCALTALED REFPERCVRGIKSREMLYNGLKFKFDIKVYSEANYLLVELKTMKAKFCESLLKRG VIVRDCYSFDGLGDNVYRVISGTFFEEVERFLKILEIIS"  
3645..4427  
/gene="MJ0956"  
3645..4427  
/gene="MJ0956"  
/note="hypothetical protein; identified by GeneMark; putative"  
/codon\_start=1  
/transl\_table=11  
/product="M. jannaschii predicted coding region MJ0956"  
/protein\_id="AAB98969.1"  
/db\_xref="GI:1499793"  
TRANSLATION="MHEDVMSLIKETHKHTILYSHGMKNRVLLGYLLKRGKYNFDE AVTIFSDRSPKWRREIKWIMKANDVHRFTSCGCKISLENNGHYDHSQVWEGIK NFEQGVASSIGEHIKDKERLUTWCNDYNLPPFHLNVRRLVKEFSLDKNLISS CYHACVFCPIYKLGSLKISIEIFYKKILITDNERKEHALNEFYDAMKEQDIDGRMINS VKIYNKYIRLELDEPYDIKSRDVEPMPDKLVNCKEYLGVK"  
4427..5479  
/gene="MJ0957"  
4427..5479  
/gene="MJ0957"  
/note="hypothetical protein; identified by GeneMark; putative"  
/codon\_start=1  
/transl\_table=11  
/product="M. jannaschii predicted coding region MJ0957"  
/protein\_id="AAB98970.1"  
/db\_xref="GI:1499794"  
TRANSLATION="MGDDIYIPPKPKPKPKPKPKYKVALWPKYTWYKFRNGTGD LLNRGNVATALEKDGSTPESGNLREDKDGKEGFEINSELSGIKYIAQIPYVK RGMLEKLNASKDKKQMEYYINHIYLVRTLEYWKGDVGVIESCDNLVGFYWNFS PGQSWGFTDWEIAQLSTIYIKQSNELNRLKLEFIWPIYINDPNNDIKRLSKYF DYVFPQHYIYIAWKYWCLNVEKNVEDIREYKYQINRYNGLTGQITGLIEVLAWI KEIPNGYIEMEVDNKIDIEKYVHDLINKACDVIKAREFLTGRDIQIRAYVFDNIENV DKVRNGAYGIKCKNW"  
5484..6764  
/gene="MJ0958"  
5484..6764  
/gene="MJ0958"  
/note="hypothetical protein; identified by GeneMark; putative"  
/codon\_start=1  
/transl\_table=11  
/product="M. jannaschii predicted coding region MJ0958"  
/protein\_id="AAB98971.1"  
/db\_xref="GI:1499795"  
TRANSLATION="MKKFLPLIILLFSIYELNVSAAEVCPCPKDGIILYHDIGYDKL FGYNDSIELIYFNKNLMDITPSDYFFPELYHNITYFECGSTNNLTLSFGYFYFP KNGDGVFLGVLYTTNNKINYTEKILWADYFVSDVDFDI SPACSPNEALLVYCF KADYSPNLIYLNNTITELKGDDEDDYTFQHYIYQPIYQPIFTYDTSKAKFYLD MRISNTSPFIYSYGGKIHPEDNITLPKYENISWECMDFYSLNGTYIYVKKLNCNII NYAGWEDYNQRPYLLIWNKTKIKISNCSPNYPYVKIKGGEIEFKYKIKFQNK YBSIRITDLYNNGILLIETNENKHLWYIYVKNNSIEETKLNKIIKLYKKHSLWDII KKLEPKIYWNHNYIIVIIAGLLWMAILWKY"

gene complement(6788.7945)  
/gene="MJ0959"  
CDS complement(6788.7945)  
/gene="MJ0959"  
/note="similar to PID:704449 GB:AE000666 percent identity:  
54.69; identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="aspartate aminotransferase (aspC)"  
/protein\_id="AAB98961.1"  
/db\_xref="GI:1591623"  
/translation="MKIDAVKLLMPGTVPVPEVLNMAIPVIGHRTKDYNNLLED  
TIEKKVFIITENDTFLITGSTAAMDMAISNIIRGDKVNLNITVNGERFANIYKA  
YKGAIRLDVWGMAEPAVKELDKYDDIKAVTVHNETSTGARNPIKRIGEVVKD  
YDALYIVDTVSLGDDYVNVKFDHIDICVTGSKCLAAPGLAAITVSEKAMEVIKKN  
DKKGVFLDLAYKKYBEKKOTPTVTPSVNLTVALNVDLDVLEEGENKRVKHEBLA  
KATRAGLEAMGIELPFAKRAEASVTTSKYPPEGIEDSKERGIILSNKYNIVVAGQKHL  
AGKPRIGHMGICGKEVLIATACVELALKELGFEVRESGVEVAKEVLLKE"  
complement(7990.8643)  
/gene="MJ0960"  
CDS complement(7990.8643)  
/gene="MJ0960"  
/note="similar to SP:P19669 PID:853766 GB:AL009126 percent  
identity: 59.02; identified by sequence similarity;  
putative"  
/codon\_start=1  
/transl\_table=11  
/product="transaldolase"  
/protein\_id="AAB98962.1"  
/db\_xref="GI:1591624"  
/translation="MKPFIDVANBEIKKYABLGVDGVTNPTLVAKEGRDFYEVVK  
BICRIVEGPAEAVISTDAEGVMVKEARELAKUADNIVIKIPMTCKGMKAVKILSARGI  
KTNVTLVPSPLQAIYAAAGATYVSPFVGRLLDIGHVGKMLIEDVYKIKYNDIKTEV  
IVASVRHPVHVEAAKIGADIATMPEFVMDKLFNRPFLTDIGLERFLKDWELKSRK"  
8860.11148  
/gene="MJ0961"  
CDS 8860.11148  
/gene="MJ0961"  
/note="similar to SP:P30665 PID:1314093 PID:3482

Query Match 27.5%; Score 29.4; DB 1; Length 17492;  
Best Local Similarity 24.8%; Pred. No. 3.5e+02;  
Matches 26; Conservative 10; Mismatches 69; Indels 0; Gaps 0;  
Qy 3 AAAAUNNNNNUNAUAGNNNNNNCUNNUUNNNNNNAAAAAUAACAUNNNNCU 62  
Db 15270 AAAATTAGTATTAGAGCTATAAATTTTAAATGCCAAAAATAAACAATTCTGTG 15329  
Qy 63 NAGNNNNNNNNNNNAGAAUNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 107  
Db 15330 AAATAAGACAGGCAAGTATTCTTAGTAGGACGAGCGGGAG 15374

RESULT 8  
AR271569\_05/c  
WPCOMMENT  
Sequence split into 17 fragments LOCUS AR271569 Accession AR271569  
Fragment Name Begin End  
AR271569\_00 1 110000  
AR271569\_01 100001 210000  
AR271569\_02 200001 310000  
AR271569\_03 300001 410000  
AR271569\_04 400001 510000  
AR271569\_05 500001 610000  
AR271569\_06 600001 710000  
AR271569\_07 700001 810000  
AR271569\_08 800001 910000  
AR271569\_09 900001 1010000  
AR271569\_10 1000001 1110000  
AR271569\_11 1100001 1210000  
AR271569\_12 1200001 1310000  
AR271569\_13 1300001 1410000  
AR271569\_14 1400001 1510000

AR271569\_15 1500001 1610000  
AR271569\_16 1600001 1664976  
Continuation (6 of 17) of AR271569 from base 500001 (AR271569 Sequence 1 from patent US  
Query Match 27.5%; Score 29.4; DB 6; Length 110000;  
Best Local Similarity 24.8%; Pred. No. 2.1e+02;  
Matches 26; Conservative 10; Mismatches 69; Indels 0; Gaps 0;  
Qy 3 AAAAUNNNNNUNAUAGNNNNNNCUNNUUNNNNNNAAAAAUAACAUNNNNCU 62  
Db 73299 AAAATTAGTATTAGAGCTATAAATTTTAAATGCCAAAAATAAACAATTCTGTG 73240  
Qy 63 NAGNNNNNNNNNNNAGAAUNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 107  
Db 73239 AAATAAGACAGGCAAGTATTCTTAGTAGGACGAGCGGGAG 73195  
RESULT 9  
AC015173/c  
LOCUS AC015173 60471 bp DNA linear HTG 16-NOV-1999  
DEFINITION Drosophila melanogaster, \*\*\* SEQUENCING IN PROGRESS \*\*\*.  
ACCESSION AC015173  
VERSION AC015173.1 GI:6436162  
KEYWORDS HTG; HTGS\_PHASE2.  
SOURCE Drosophila melanogaster (fruit fly)  
ORGANISM Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
REFERENCE 1 (bases 1 to 60471)  
AUTHORS Adams,M. and Venter,J.C.  
TITLE Direct Submission  
JOURNAL Submitted (16-NOV-1999) Celera Genomics, 45 West Gude Drive,  
Rockville, MD, USA  
COMMENT This sequence was identified as CDM:10210450 by the submitter.  
For further information on this sequence e-mail to fly@celera.com.  
\* NOTE: This is a 'working draft' sequence.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.

FEATURES  
source  
1.60471  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"  
ORIGIN  
Query Match 27.3%; Score 29.2; DB 2; Length 60471;  
Best Local Similarity 25.0%; Pred. No. 2.7e+02;  
Matches 21; Conservative 13; Mismatches 50; Indels 0; Gaps 0;  
Qy 1 GGAAAAUNNNNNUNAUAGNNNNNNCUNNUUNNNNNNAAAAAUAACAUNNNNC 60  
Db 15041 GGAAAAATAAATCTAGTATTACAAATCTTAGTTTGTACCAAAAAATAAATCTGTAA 14982  
Qy 61 UUNAGNN 84  
Db 14981 TTAATTTTAAATTCAAATTTAAAT 14958

RESULT 10  
AC009846/c  
LOCUS AC009846 164038 bp DNA linear INV 17-FEB-2001  
DEFINITION Drosophila melanogaster, chromosome 3R, region 94C-94D, BAC clone  
BACR23P10, complete sequence.  
ACCESSION AC009846  
VERSION AC009846.9 GI:12957616  
KEYWORDS HTG.  
SOURCE Drosophila melanogaster (fruit fly)  
ORGANISM Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.



```

Rubin,G.M., Mungall,C.J. and Lewis,S.E.
Annotation of Drosophila melanogaster genome
Unpublished
4 (bases 1 to 234378)
Adams,M.D., Celniker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,C.J.
Direct Submission
Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
5 (bases 1 to 234378)
FlyBase
Direct Submission
Submitted (06-SEP-2002) University of California Berkeley, 539 Life
Sciences Addition, Berkeley, CA 94720, USA
6 (bases 1 to 234378)
FlyBase
Direct Submission
Submitted (13-FEB-2003) University of California Berkeley, 539 Life
Sciences Addition, Berkeley, CA 94720, USA
Sequence update by submitter
On Sep 18, 2002 this sequence version replaced gi:10726705.
Location/Qualifiers
1..234378
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/chromosome="3R"
/note="genotype: Y[1]; cn[1] bw[1] sp[1]; Rh6[1]"
10750..13380
 locus_tag="CG17622"
/note="last curated on Tue Jan 22 00:31:41 PST 2002"
/map="94B10-94B10"
/db_xref="FLYBASE:FBgn0038998"
join(10750..10772,11010..11215,12157..12598,12859..13185,
13311..13380)
 locus_tag="CG17622"
/product="CG17622-PA"
/db_xref="FLYBASE:FBgn0038998"
join(10750..10772,11010..11215,12157..12598,12859..13185,
13311..13380)
 locus_tag="CG17622"
/note="CG17622 gene product"
/codon_start=1
/product="CG17622-PA"
/protein_id="AAF56036.1"
/db_xref="GI:7300896"
/db_xref="FLYBASE:FBgn0038998"
/translaton="MVPFFRSVAQSAASAAAGAGGAGGGGGSGSQRPGSRKL
PPLLSRTLPALIVGLPVYLFHTDKQTQLDERLIGSKRSGSAGSHWSLLTRAT
NSSSMNNPPASSSSSSNNNNNHNANNNTIMVKSLLPKDDSNLVYRRKSSGST
PHOQHARRHHQQHHLQOOQOOQOOLOHLDGSLFOLGDDFECHAD
GSLIRLPAPHVVAARVRLAARKRAASEASTPMREHTLHEMPLAESACSGSATG
SGTIFTRLAKLLQSSTLAGRSNLQPPMDNAVQSLGRGLGLGREAPRRLSWESW
KADLRVSLAAPNGRPPEVG"
18885..19523
 locus_tag="CG17623"
/note="last curated on Tue Jan 22 00:33:38 PST 2002"
/map="94B11-94B11"
/db_xref="FLYBASE:FBgn0038999"
join(18885..18931,19007..19523)
 locus_tag="CG17623"
/product="CG17623-RA"
/db_xref="FLYBASE:FBgn0038999"
join(18885..18931,19007..19523)
 locus_tag="CG17623"
/note="CG17623 gene product"
/codon_start=1
/product="CG17623-PA"
/protein_id="AAF56037.2"
/db_xref="GI:23171989"
/db_xref="FLYBASE:FBgn0038999"
/translat.on="MPSALHLRESFPESIERKDSALQRSIDSFAEIVFSESPRFS

```

LVTGPGACASGASGSPFKRPSASSLYSTSTSSSFASQMAQLNNVGDLSAGSGS  
 ATGCGSHHGGAGSGGNGGASGSSRRSMSPSSTRSKLTRIINGECGVNRMRNP  
 KIYQANRALCMDSTSIQTSLYLYLLTI"  
 28305..36625  
 /locus\_tag="CG6954"  
 /note="last curated on Tue Jan 22 00:49:10 PST 2002"  
 /map="94C1-94C1"  
 /db\_xref="FLYBASE:FBgn0039000"  
 join(28305..28975,30084..30334,31045..31599,31729..31807,  
 31872..31975,32036..32173,32240..32464,32702..32908,  
 32966..33176,33240..36625)  
 /locus\_tag="CG6954"  
 /product="CG6954-PA"  
 /db\_xref="FLYBASE:FBgn0039000"  
 join(28305..28975,30084..30334,31045..31599,31729..31807,  
 31872..31975,32036..32173,32240..32464,32702..32908,  
 32966..33176,33240..35715)  
 /locus\_tag="CG6954"  
 /note="CG6954 gene product"  
 /codon\_start=1  
 /product="CG6954-PA"  
 /protein\_id="AAF56038.1"  
 /db\_xref="GI:7300898"  
 /db\_xref="FLYBASE:FBgn0039000"  
 /translation="MERKALSSYAKFKDQBAAAAVTPQKONESDDSLPALRRSHSV  
 RSSLRRLKNLHSPALWSQSPKLTQLHMRNRSRGSTLDKSKAARLMLGLDIDA  
 VAVLITVKYKNEFVPLKRAMNDSCDHDELGIERLMSSLPYGAKLPRACKLQIPE  
 SYCQAEVAPKATLDLEKTSFGBESSAVLADIITHAQQGIYGTMTRTAKKPE  
 YLNGRCYHSDSTKLPMNPKLPMSCLRSPAFVRQKFTLNETRFLCQKSPHR  
 STATATNSPSSPAAQQORHSHVADINMLPKRDGDISVRLREKRVNPSG  
 GNNNMNDVPTYPAPLLSTSRWQLTQHAAVAKIRTAISCHQDLREMEFLPHSR  
 TDSQNSQOJLCISHSOPKAKLRDHDHDTGSSLDSPGLDENASANGQLSVHROP  
 VQLSKIPGNTPHRASKSKQLAKQQLKLTQNNLHLHALFSVHGHLEKART  
 ILESTDVDSINNDGALDIADVLSNNRMTMLQHGAVESQSPSDITGKINGL  
 LKDAESRTHDLSGPEGLCPVPFASRPSISIIIGNSSAVTGCTSEVEKQIGIWERR  
 VKGLRLQDWDQARPPAPASVVDVTGDSISVQILPFGAIGTKFKVQWSTEAD  
 FNNVVEGELLEWISPHGTMAQCHISGLTQGRRYFLRAACGNVKGWGTYRTSVPASV  
 VPSIWRDLNDRDVRVGBRLLDNLFTAVRLPADVSELTLDPASQORNPKKTTI  
 MKVACTDVEDKSLDMERNLTSVAHFRTKLLSAVCQMSAUGITDGLQYIKPLRDA  
 QGVTLTCQSVKSKQVSLNSRVPFKLQKLGHALDEYINELLSSIGDQLHY  
 QQAALORLEPLGLVLYKMGCSIDQIVVVPKTNVLPCHKVREYTPAEWQVLIH  
 RYSDPLDPLDFSAGDGGSGATATEVQRLFLYDLTNAMHKLFSAMNKKADAT  
 HRLYDVEIENSHPDLISFLVVCPSASSCAVPGQSELLQRDDLSLISIOAFEMHLRT  
 YOPALIQKRYASCEILDALATHSLRAFSSSELQAKERLQELASLTIWVK  
 SVRWLMVYVAYRNKAQPSLAWREILDFAQQRDEAVATSAGSANGQLQLLPFIRS  
 KFKSTQGRGKWPFGPSTSDQSNPEHKSQNLLESAITPYEPASNQVPTQQQPPQ  
 QSQSQQLLOVNSISEYAGSICSEVFRKNSGMSSTYTSRFSYSAVDSASDGNSTN  
 SVFATPPRSDDTLADALPHSOAAQAKRTNSIAHTNPLITVHSSSAPYLAGSS  
 VSLRSGNGAFATDLEHKPLKPAITKANSSNREGGPKYLSKLAEALRVAGVEEPAVSTS

Query Match 27.3%; Score 29.2; DB 3; Length 234378;  
 Best Local Similarity 25.0%; Pred. No. 1.9e+02;  
 Matches 21; Conservative 13; Mismatches 50; Indels 0; Gaps 0;  
 QY 1 GGAAAAUNNNNNNUAUGNNNNNNCUNNNUNNNNNNAAAAAANUANAACAUNNNNC 60  
 Ddb 101887 GGAAAAATAAACTTAGTATTACAACTTAGTTTGTACCAAAAAATAAAATCTGTATA 101828  
 QY 61 UUNAGNNNNNNNNNNNNNAGAAU 84  
 Ddb 101827 TTAATTTTAAATTCAAATTTAAAT 101804  
 RESULT 12  
 AC079951 191942 bp DNA linear PRI 15-MAR-2002  
 LOCUS Homo sapiens 12q BAC RP11-346L1 (Roswell Park Cancer Institute  
 DEFINITION Human BAC Library) complete sequence.  
 AC079951  
 ACCESSION AC079951  
 VERSION AC079951.35 GI:15055247  
 KEYWORDS HTG.  
 SOURCE Homo sapiens (human)

```
Query Match      27.3%; Score 29.2; DB 3; Length 234378;
Best Local Similarity 25.0%; Pred.No. 1.9e+02; Indels    0; Gaps    0;
Matches 21; Conservative 13; Mismatches 50;

1 GGAAAAUUNNNNUNAUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUAAACAUNNNNC   60
||||| : : : : | : |||| : ||| :
101887 GGAATAAAGCTTAGTATTTCACATCTTAGTTTGTAACCAAAAAATAAAAACTGTATA   101828
                                     : : : : :

61 UUNAGNNNNNNNNNNNNNAGAUAU 84
   :: |||
101827 TTAATTTTAAATTCAAATTTAAAT 101804

RESULT 12
AC079951     LOCUS              AC079951     DNA       linear   PRI 15-MAR-2002
DEFINITION   Homo sapiens 12q BAC RP11-346L11 (Roswell Park Cancer Institute
             Human BAC Library) complete sequence.
ACCESSION   AC079951
VERSION      AC079951.35 GI:15055247
KEYWORDS     HTG.
SOURCE       Homo sapiens (human)
```

## ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

## AUTHORS

1 (bases 1 to 191942)  
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Albrooks,S.L., Anaratunge,H.C., Are,J.R., Banks,T., Barbaria,J., Benton,J., Binaige,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.P., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Emerling,S., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Han,J., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homsif,F., Howard,S., Huber,J., Huiyk,S., Hume,J., Ioshikhes,I., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lee,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lied,C., Liu,J., Liu,W., Louissegh,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Marondel,I., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Merscher,S., Metzker,M., Miller,A., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Montgomery,K.T., Morgan,M., Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogih,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peary,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.B., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savary,G., Scherer,S., Scott,G., Shen,H., Shim,C., Shoostari,N., Sisson,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williams,A., Wleczky,R., Woodson,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Kucherlapati,R., Weinstock,G. and Gibbs,R.

Direct Submission  
Unpublished  
2 (bases 1 to 191942)  
Worley,K.C.  
Direct Submission  
Submitted (20-SEP-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 191942)  
Worley,K.C.  
Direct Submission  
Submitted (01-AUG-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
4 (bases 1 to 191942)  
Worley,K.C.  
Direct Submission  
Submitted (09-AUG-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
5 (bases 1 to 191942)  
Worley,K.C.  
Direct Submission  
Submitted (15-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Aug 1, 2001 this sequence version replaced gi:14961675.  
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email

gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

## ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:  
<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

## QUALSTAT-REPORT.

FEATURES	source	Location/Qualifiers
repeat_region	582..883	/rpt_family="AluSg"
repeat_region	1600..1799	/rpt_family="MER46A"
repeat_region	1906..2212	/rpt_family="AluSg"
repeat_region	3606..3915	/rpt_family="AluSg"
repeat_region	4874..4917	/rpt_family="AluSc"
repeat_region	complement(5619..5883)	/rpt_family="AluSg"
repeat_region	6028..6335	/rpt_family="AluSx"
repeat_region	6341..6774	/rpt_family="L1MCb"
repeat_region	6318..7034	/rpt_family="L1M4"
repeat_region	7038..7264	/rpt_family="L1MA2"
repeat_region	complement(7337..7633)	/rpt_family="HUES-P3b"
repeat_region	complement(7659..11980)	/rpt_family="L1PA5"
STS	9027..9204	/standard_name="61535"
repeat_region	complement(11981..13272)	/rpt_family="SVA"
repeat_region	complement(13273..15103)	







Moule, S., Mungall, K., Murphy, L., Oliver, K., Quail, M.A.,  
 Rajandream, M.-A., Rutter, S., Skelton, J., Squares, R., Squares, S.,  
 Sulston, J.E., Whitehead, S., Woodward, J.R., Newbold, C. and  
 Barrell, B.G.  
 The complete nucleotide sequence of chromosome 3 of *Plasmodium*  
*falciparum*  
 Nature 400 (6744), 532-538 (1999)  
 99376085  
 10449855  
 2  
 Hall, N., Pain, A., Berriman, M., Churcher, C., Harris, B., Harris, D.,  
 Mungall, K., Bowman, S., Atkin, R., Baker, S., Barron, A., Brooks, K.,  
 Buckee, C.O., Burrows, C., Cherevach, I., Chillingworth, C.,  
 Chillingworth, T., Christodoulou, Z., Clark, L., Clark, R., Corton, C.,  
 Cronin, A., Davies, R., Davis, P., Dear, P., Dearden, F., Doggett, J.,  
 Fellwell, T., Goble, A., Goodhead, I., Gwilliam, R., Hamlin, N.,  
 Hance, Z., Harper, D., Hauser, H., Hornsby, T., Holroyd, S.,  
 Horrocks, P., Humphray, S., Jagels, K., James, K.D., Johnson, D.,  
 Kerhornou, A., Knights, A., Konfortov, B., Kyes, S., Larkin, N.,  
 Lawson, D., Lennard, N., Line, A., Maddison, M., McLean, J., Mooney, P.,  
 Moule, S., Murphy, L., Oliver, K., Ormond, D., Price, C., Quail, M.A.,  
 Rabinowitsch, E., Rajandream, M.A., Rutter, S., Rutherford, K.M.,  
 Sanders, M., Simmonds, M., Seeger, K., Sharp, S., Smith, R., Squares, R.,  
 Squares, S., Stevens, K., Taylor, K., Tivey, A., Unwin, L.,  
 Whitehead, S., Woodward, J., Sulston, J.E., Craig, A., Newbold, C. and  
 Barrell, B.G.  
 Sequence of *Plasmodium falciparum* chromosomes 1, 3-9 and 13  
 Nature 419 (6906), 527-531 (2002)  
 22255708  
 12368867  
 3 (bases 1 to 113880)  
 Lawson, D., Bowman, S. and Barrell, B.  
 Direct Submission  
 Submitted (18-NOV-1998) *P.falciparum* Genome Sequencing Consortium,  
 The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge  
 CB10 1SA, UK  
 On Apr 30, 2000 this sequence version replaced gi:4493924.  
 For more information about this sequence or the Malaria Project,  
 see [http://www.sanger.ac.uk/Projects/P\\_falciparum](http://www.sanger.ac.uk/Projects/P_falciparum).

FEATURES  
source

```

/organism="Plasmodium falciparum 3D7"
/mol_type="genomic DNA"
/isolate="3D7"
/db_xref="taxon:36329"
/chromosome="3"
/clone="MAL3P4"
2. .1024
/gene="MAL3P4.27"
/note="synonym: PFC0430w"
<2. .1024
/gene="MAL3P4.27"
/note="Revised: removed 5' exon"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAD49086.1"
/db_xref="GI:23476984"
/db_xref="SPTREMBL:O81222"
/translations="KNVDSIQDLIKSIDFWILKENIKMIDIPFQKYLIKVKNIILK
NEMDVCVDYIVGCIENEFFMKNEHKQNGYLKSEIKYKNNIDNDFECCKEVINENFI
EEFTSDSPNDLIDKNEKDISDNNKNNKNNNNNNNNNNNNDDDFNNNSYIQNN
FLEENIRKKEFIPYKDNNSNIKIQNTQYDKQNKSFQENINILKSPQSQN
LSVMFDRNYFHCINKYIKSKNKEIGLIKKLLTYLDLRSYNTHGLNKLNTFORLM
HIEKIDNNTKG:GETLNTVSCDINPVLNNTSYVGEGPKKYLIYSIPGKKIKKNI
KKNKIK"

```

misc feature

/notes:"Possible exon as peptide fragment matches one of the orphan peptides from proteomics study [source: M. Mann]

[illegible]





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

MM nucleic - nucleic search, using sw model  
Run on: March 8, 2004, 03:12:12 ; Search time 307 Seconds  
(without alignments)  
1480.642 Million cell updates/sec

Title: US-09-646-651C-2  
Perfect score: 107  
Sequence: 1 ggaagaaunnnnnunauaugh.....nnnnnnnnnnuuuagcag 107

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : N Geneseq\_29Jan04.\*
- 1: Geneseqn1980s.\*
  - 2: Geneseqn1990s.\*
  - 3: Geneseqn2000s.\*
  - 4: Geneseqn2001as.\*
  - 5: Geneseqn2001bs.\*
  - 6: Geneseqn2002s.\*
  - 7: Geneseqn2003as.\*
  - 8: Geneseqn2003bs.\*
  - 9: Geneseqn2003cs.\*
  - 10: Geneseqn2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	31.4	29.3	298	Aax23773	Aax23773 Ribonucle
2	29.4	27.5	110000	AAV21209_05	Continuation (6 of
3	29.2	27.3	3456	ABL18738	Ab118738 Drosophil
4	29.2	27.3	4829	ABL18730	Ab118730 Drosophil
5	27.4	25.6	524	ABQ89024	Abq89024 Human pro
6	27.4	25.6	524	ADB81968	ADB81968 Human cDN
7	27.4	25.6	761	ADB82181	ADB82181 Human cDN
8	27.4	25.6	769	ABQ89163	Abq89163 Human pro
9	27.4	25.6	769	ADB82106	ADB82106 Human cDN
10	27.4	25.6	789	ABQ89160	Abq89160 Human pro
11	27.4	25.6	789	ADB82103	ADB82103 Human cDN
12	27.4	25.6	835	ABQ89238	Abq89238 Human pro
13	27.4	25.6	929	ADB83094	ADB83094 Human cDN
14	27.4	25.2	325791	AAAS43104	AAAS43104 Human Oes
15	26.8	25.0	6419	ABL32266	ABL32266 Human imm
16	26.8	25.0	13814	ABL33193	ABL33193 Human imm
17	26.6	24.9	10034	ABL33885	ABL33885 Human imm
18	26.2	24.5	100543	ABSS2816	ABSS2816 Genomic D
19	26	24.3	318	ADA29355	Ada29355 DNA encod
20	26	24.3	5504	ABL70350	Ab170350 Chemical
21	26	24.3	5504	AAAS61309	AAAS61309 Human gen
22	26	24.3	7977	Ab134089	Ab134089 Human imm
23	25.8	24.1	250	AAAC16119	AAAC16119 Human sec

ALIGNMENTS

RESULT 1	
AAX23773	
ID AAX23773 standard; RNA; 298 BP.	
XX	
AC AAX23773;	
XX	
DT 21-JUN-1999 (first entry)	
XX	
DE Ribonucleotide protein ARNA VI Clone P10 RNA.	
XX	
KW Angiotropin related protein; ARP; ternary complex; S100 protein;	
KW copper-containing ribonucleoprotein; copper; cell selective;	
KW morphogenic action; blood capillary endothelial cell; confluent;	
KW non-mitogenic induction; cell phenotype; three-dimensional organoid;	
KW spatiotemporal supracellular organisation; chemotropic; blood vessel;	
XX tissue neovascularisation; angiogenesis modulation; ss.	
OS Synthetic.	
XX	

```

confluent state, for non-mitogenic alteration of the spatiotemporal
supracellular organisation of cells into three-dimensional organoid,
capillary-like structures in culture, (2) a specific chemotropic action
on blood vessels in vivo, (3) induction of directional growth of blood
vessels in vivo and (4) induction of neovascularisation of tissues
through directed ingrowth of blood vessels. Their use for modulating
angiogenesis is claimed
XX
SQ      Sequence 298 BP; 103 A; 40 C; 38 G; 0 T; 117 U; 0 Other;

      Query Match      29.3%; Score 31.4; DB 2; Length 298;
      Best Local Similarity 42.1%; Pred. No. 6.2;
      Matches 32; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY      32 UUNNNNNAAAAAUAAPAAACUUNNNNNCUUAGNNNNNNNNNNNNNNAGAAUUNNNNN 91
Db      89 UUUGUACUAUAAAUAAAAACAGUCUUCUUUAGACACUCCUUAUAGAAUAAAUAA 148

QY      92 NNNNNNNNNUUAGCAG 107
Db      149 AUAAACUAUAGCAG 164

RESULT 2
AAV21209_05/c
Continuation (6 of 17) of AAV21209 from base 500001 (Methanococcus jannaschii circular
WP Sequence split into 17 fragments LOCUS AAV21209 Accession Aav21209
WP Fragment Name Begin End
WP AAV21209_01 1 110000
WP AAV21209_01 100001 210000
WP AAV21209_02 200001 310000
WP AAV21209_03 300001 410000
WP AAV21209_04 400001 510000
WP AAV21209_05 500001 610000
WP AAV21209_06 600001 710000
WP AAV21209_07 700001 810000
WP AAV21209_08 800001 910000
WP AAV21209_09 900001 1010000
WP AAV21209_10 1000001 1110000
WP AAV21209_11 1100001 1210000
WP AAV21209_12 1200001 1310000
WP AAV21209_13 1300001 1410000
WP AAV21209_14 1400001 1510000
WP AAV21209_15 1500001 1610000
WP AAV21209_16 1600001 1664976

      Query Match      27.5%; Score 29.4; DB 2; Length 110000;
      Best Local Similarity 24.8%; Pred. No. 47;
      Matches 26; Conservative 10; Mismatches 69; Indels 0; Gaps 0;

QY      3 AAAAAUUNNNNUAUAUGUNNNNNNUNUUNNNNNNNAAAAAUAUAAACAUUNNNNCUU 62
Db      73299 AAAATTAGTATTAGAGCTATACTATTATTAAATGCCAAAAAATAAACATTTCTGGTG 73240

QY      63 NAGNNNNNNNNNNNAGAAUUNNNNNNNNNNNNNNNNNNNNNNNNUAGCAG 107
Db      73239 AAATAATGACAGGCAAAAGTATCTTTAGTAGCAGACCGGGAG 73195

RESULT 3
ABL18738
ID ABL18738 standard; DNA; 3456 BP.
XX
XX ABL18738;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 7687.
DE
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
XX Drosophila melanogaster.
OS

```

```

XX PN WC200171042-A2.
XX XX
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US009231.
XX PR 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX PA (PEKE ) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX DR WPI; 2001-656860/75.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signaling and cell-cell
XX PT interactions.
XX PS Claim 1; SEQ ID NO 7687; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
XX CC ABB72072). The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 3456 BP; 827 A; 939 C; 888 G; 802 T; 0 U; 0 Other;

Query Match 27.3%; Score 29.2; DB 4; Length 3456;
Best Local Similarity 25.0%; Pred. No. 34;
Matches 21; Conservative 13; Mismatches 50; Indels 0; Gaps 0

Qy 1 GGAAAAUNNNUNUAUGNNNNNNNCUNNUUNNNNNNNAAAAAUAACAUNNNNC 60
Db 769 GGAAAAATAAACTTAGTATTACAATCTTAGTTTGACCAAAAAATAAAATCTGTATA 828
Qy 61 UUNAGNNNNNNNNNNNAGAAU 84
Db 829 TTAATTTTAAATTCAAATTTAAAT 852

RESULT 4
ABL18730
ID ABL18730 standard; DNA; 4829 BP.
XX AC ABL18730;
XX AC
XX XX
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 7663.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical; gene; ds.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US009231.
XX PR 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX PA (PEKE ) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX DR WPI; 2001-656860/75.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signaling and cell-cell
XX PT interactions.
XX PS Claim 1; SEQ ID NO 7687; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
XX CC ABB72072). The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 3456 BP; 827 A; 939 C; 888 G; 802 T; 0 U; 0 Other;

```









[illegible]





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

DM nucleic - nucleic search, using sw model

Run on: March 8, 2004, 04:31:47 ; Search time 60 seconds  
(without alignments)  
989.662 Million cell updates/sec

Title: US-09-646-651C-2

Perfect score: 107

Sequence: 1 ggaaauunnnunauaugh.....nnnnnnnnnnuuuagcag 107

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:\*

- 1: /cgn2\_6/prodata/2/ina/5A COMB.seq:\*
- 2: /cgn2\_6/prodata/2/ina/5B COMB.seq:\*
- 3: /cgn2\_6/prodata/2/ina/6A COMB.seq:\*
- 4: /cgn2\_6/prodata/2/ina/6B COMB.seq:\*
- 5: /cgn2\_6/prodata/2/ina/pOTUS COMB.seq:\*
- 6: /cgn2\_6/prodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	29.4	27.5	1664976	4	US-08-916-421B-1
C 2	26	24.3	318	4	US-09-328-352-642
C 3	25.8	24.1	934	3	US-08-842-306B-5
C 4	25.8	24.1	934	3	US-08-838-973B-5
C 5	25.8	24.1	934	4	US-08-771-212A-5
C 6	24.6	23.0	499	4	US-09-621-976-13988
C 7	24.6	23.0	3552	4	US-09-643-597-126
C 8	24.6	23.0	3552	4	US-09-480-884A-126
C 9	24.6	23.0	3552	4	US-09-542-615A-126
C 10	24.6	23.0	3552	4	US-09-606-421B-126
C 11	24.6	23.0	3552	4	US-09-221-107-126
C 12	24.6	23.0	5652	4	US-09-601-198-75
C 13	24.4	22.8	398	4	US-09-401-064-281
C 14	24.4	22.8	5614	4	US-08-956-171B-99
C 15	24.4	22.8	5915	4	US-10-204-708-77
C 16	24.2	22.6	8093	4	US-10-204-708-32
C 17	24.2	22.6	640681	4	US-09-790-988-1
C 18	24	22.4	927	4	US-09-134-000C-3089
C 19	24	22.4	6583	4	US-10-204-708-25
C 20	24	22.4	11131	4	US-10-204-708-28
C 21	23.8	22.2	65792	4	US-09-596-002-31
C 22	23.6	22.1	1474	4	US-08-858-207A-144
C 23	23.6	22.1	2418	4	US-09-601-198-61
C 24	23.6	22.1	8703	4	US-08-961-527-177
C 25	23.4	21.9	2319	2	US-08-966-389-5
C 26	23.4	21.9	2319	2	US-09-103-509-5
C 27	23.4	21.9	2319	2	US-09-102-644-5

28	23.4	21.9	2319	2	US-09-218-032-5	Sequence 5, Appli
C 29	23.4	21.9	2675	4	US-09-636-791A-12	Sequence 12, Appli
C 30	23.4	21.9	3345	2	US-08-966-389-8	Sequence 8, Appli
C 31	23.4	21.9	3345	2	US-09-103-509-8	Sequence 8, Appli
C 32	23.4	21.9	3345	2	US-09-102-644-8	Sequence 8, Appli
C 33	23.4	21.9	3345	2	US-09-218-032-8	Sequence 8, Appli
C 34	23.4	21.9	9347	4	US-10-204-708-35	Sequence 35, Appli
C 35	23.4	21.9	1664976	4	US-08-916-421B-1	Sequence 1, Appli
C 36	23.2	21.7	711	1	US-08-307-499-49	Sequence 49, Appli
C 37	23.2	21.7	711	3	US-09-299-268-49	Sequence 49, Appli
C 38	23.2	21.7	727	1	US-07-885-970A-6	Sequence 6, Appli
C 39	23.2	21.7	727	1	US-08-298-687A-6	Sequence 6, Appli
C 40	23.2	21.7	727	1	US-08-530-797-5	Sequence 5, Appli
C 41	23.2	21.7	727	1	US-08-298-829-6	Sequence 6, Appli
C 42	23.2	21.7	727	2	US-08-787-335-5	Sequence 5, Appli
C 43	23.2	21.7	1128	4	US-09-543-681A-3553	Sequence 3553, Ap
C 44	23.2	21.7	10144	4	US-10-204-708-94	Sequence 94, Appli
C 45	23.2	21.7	14176	1	US-08-307-499-1	Sequence 1, Appli

#### ALIGNMENTS

#### RESULT 1

US-08-916-421B-1/c

; Sequence 1, Application US/08916421B

; Patent No. 6503729

; GENERAL INFORMATION:

; APPLICANT: Bult et al.

; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococ

; Patent No. 6503729

; FILE REFERENCE: PB275

; CURRENT APPLICATION NUMBER: US/08/916,421B

; PRIOR FILING DATE: 1997-08-22

; PRIOR APPLICATION NUMBER: US 60/024,428

; PRIOR FILING DATE: 1996-08-22

; NUMBER OF SEQ ID NOS: 3

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1

; LENGTH: 1664976

; TYPE: DNA

; ORGANISM: Methanococcus jannaschii

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (28222)..(28222)

; OTHER INFORMATION: n equals a, t, c, or g

; NAME/KEY: misc feature

; LOCATION: (28257)..(28258)

; OTHER INFORMATION: n equals a, t, c, or g

; NAME/KEY: misc feature

; LOCATION: (84773)..(84773)

; OTHER INFORMATION: n equals a, t, c, or g

; NAME/KEY: misc feature

; LOCATION: (84808)..(84808)

; OTHER INFORMATION: n equals a, t, c, or g

; NAME/KEY: misc feature

; LOCATION: (84812)..(84812)

; OTHER INFORMATION: n equals a, t, c, or g

; NAME/KEY: misc feature

; LOCATION: (98120)..(98120)

; OTHER INFORMATION: n equals a, t, c, or g

; NAME/KEY: misc feature

; LOCATION: (98159)..(98159)

; OTHER INFORMATION: n equals a, t, c, or g

; NAME/KEY: misc feature

; LOCATION: (98239)..(98239)

; OTHER INFORMATION: n equals a, t, c, or g

; NAME/KEY: misc feature

; LOCATION: (98266)..(98266)

; OTHER INFORMATION: n equals a, t, c, or g

; NAME/KEY: misc feature

; LOCATION: (98343)..(98343)







```

Query Match          23.0%; Score 24.6; DB 4; Length 3552;
Best Local Similarity 27.2%; Pred.No.38;
Matches      22; Conservative    8; Mismatches   51; Indels     0; Gaps       0;

#           3 AAAAAUUNNNNUAUGNUNNNNCUNUUUINNUNNNNAAAAUAAAUAACCAUUNNNNCUU 62
|||:::~::~|||:::~::~|||||:::~::~|||||:::~::~|||||:::~::~|||||:::~::~|||:
3081 AAATCACACTATTATGTTATTCCTTAAAGTGATGTGGTTTTAAAAATAAACAGAATAATT 3140

#           63 NAGNNNNNNNNNNNNNAGAA 83
|||
3141 GAGTATCACTATCGTAGAGAA 3161


=====
RESULT 9
:-09-542-615A-126
Sequence 126, Application US/09542615A
Patent No. 6518256
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Fan, Liqun
APPLICANT: Kalos, Michael D.
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Hosken, Nancy A.
APPLICANT: Fanger, Gary R.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.455C8
CURRENT APPLICATION NUMBER: US/09/542,615A
CURRENT FILING DATE: 2000-04-14
NUMBER OF SEQ ID NOS: 350
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 126
LENGTH: 3552
TYPE: DNA
ORGANISM: Homo sapien
:-09-542-615A-126


Query Match          23.0%; Score 24.6; DB 4; Length 3552;
Best Local Similarity 27.2%; Pred.No.38;
Matches      22; Conservative    8; Mismatches   51; Indels     0; Gaps       0;

#           3 AAAAAUUNNNNUAUGNUNNNNCUNUUUINNUNNNNAAAAUAAAUAACCAUUNNNNCUU 62
|||:::~::~|||:::~::~|||||:::~::~|||||:::~::~|||||:::~::~|||||:::~::~|||:
3081 AAATCACACTATTATGTTATTCCTTAAAGTGATGTGGTTTTAAAAATAAACAGAATAATT 3140

#           63 NAGNNNNNNNNNNNNNAGAA 83
|||
3141 GAGTATCACTATCGTAGAGAA 3161


=====
RESULT 10
:-09-606-421B-126
Sequence 126, Application US/09606421B
Patent No. 6531315
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Fan, Liqun
APPLICANT: Kalos, Michael D.
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Hosken, Nancy
APPLICANT: Fanger, Gary R.
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.455C9
CURRENT APPLICATION NUMBER: US/09/606.421B
CURRENT FILING DATE: 2000-06-28
NUMBER OF SEQ ID NOS: 358
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 126
```

```

; LENGTH: 3552
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-606-421B-126

Query Match      23.0%; Score 24.6; DB 4; Length 3552;
Best Local Similarity 27.2%; Pred. No. 38;
Matches 22; Conservative 8; Mismatches 51; Indels 0; Gaps 0;

QY 3 AAAAUNNNNNUNAAUAGUNNNNNNNCUNNNUNNNNNNAAAAAUAUAAACAUNNNNNCUU 62
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3081 AATATCACATATTATGTATTCTTAACTGATAGTTTAAAAAATAACAGAAATATT 3140
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 63 NAGNNNNNNNNNNNNNAGAAA 83
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3141 GAGTATCACTATGTGAAGAA 3161
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 11
US-09-221-107-126
; Sequence 126, Application US/092211107
; Patent No. 6660838
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
; FILE REFERENCE: 210121.455C2
; CURRENT APPLICATION NUMBER: US/09/221,107
; CURRENT FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 161
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 126
; LENGTH: 3552
; TYPE: DNA
; ORGANISM: Human
US-09-221-107-126

Query Match      23.0%; Score 24.6; DB 4; Length 3552;
Best Local Similarity 27.2%; Pred. No. 38;
Matches 22; Conservative 8; Mismatches 51; Indels 0; Gaps 0;

QY 3 AAAAUNNNNNUNAAUAGUNNNNNNNCUNNNUNNNNNNAAAAAUAUAAACAUNNNNNCUU 62
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3081 AATATCACATATTATGTATTCTTAACTGATAGTTTAAAAAATAACAGAAATATT 3140
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 63 NAGNNNNNNNNNNNNNAGAAA 83
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3141 GAGTATCACTATGTGAAGAA 3161
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 12
US-09-601-198-75/c
; Sequence 75, Application US/09601198
; Patent No. 6531583
; GENERAL INFORMATION:
; APPLICANT: Cassell, Gail H.
; APPLICANT: Chen, Ellison Y.
; APPLICANT: Glass, Jennifer S.
; APPLICANT: Glass, John I.
; APPLICANT: Heiner, Cheryl R.
; APPLICANT: Lefkowitz, Elliot
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREA PLASMA
; TITLE OF INVENTION: UREA LYTICUM
; FILE REFERENCE: UAB-13452/22
; CURRENT APPLICATION NUMBER: US/09/601,198
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/073,189
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 75
; LENGTH: 5652
; TYPE: DNA
; ORGANISM: Ureaplasma urealyticum

```







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

M nucleic - nucleic search, using sw model

Run on: March 8, 2004, 04:40:28 ; Search time 1220 Seconds

(without alignments)  
320.775 Million cell updates/sec

Title: US-09-646-651C-2

Perfect score: 107

Sequence: 1 ggaaauunnnunnaugna.....nnnnnnnnnnuuuagac 107

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2421054 seqs, 1828716029 residues

Total number of hits satisfying chosen parameters: 4842108

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09C\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
18: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

result No.	Score	Query Match %	Length	ID	Description
1	27.4	25.6	524	15	US-10-012-697-280
2	27.4	25.6	761	15	Sequence 280, App
3	27.4	25.6	769	15	Sequence 493, App
4	27.4	25.6	789	15	GENERAL INFORMAT
5	27.4	25.6	929	15	Sequence 415, App
6	26.8	25.0	6419	14	Sequence 1307, App
7	26.8	25.0	13814	14	Sequence 239, App
8	26.8	25.0	3673778	14	Sequence 1166, App
9	26.6	24.9	487	15	Sequence 323248,
10	26.6	24.9	488	15	Sequence 324208,
11	26.6	24.9	634	15	Sequence 8286, App
12	26.6	24.9	10034	14	Sequence 1858, App
13	26.2	24.5	390	10	Sequence 17808, A
14	26.2	24.5	3673778	14	Sequence 1, Appli
15	26	24.3	511	12	Sequence 85840, A

C	16	26	24.3	5504	12	US-10-221-613-270	Sequence 270, App
	17	26	24.3	7977	14	US-10-311-455-2062	Sequence 2062, App
	18	26	24.3	3673778	14	US-10-312-841-1	Sequence 1, Appli
	19	25.8	24.1	553	9	US-09-560-863-928	Sequence 928, App
	20	25.8	24.1	585	15	US-10-027-632-251034	Sequence 251034,
	21	25.8	24.1	934	9	US-09-945-249-5	Sequence 5, Appli
	22	25.8	24.1	1031	14	US-10-093-524-9	Sequence 9, Appli
	23	25.8	24.1	1386	12	US-10-282-122A-15352	Sequence 15352, A
	24	25.8	24.1	1566	14	US-10-081-051-59	Sequence 59, Appl
	25	25.8	24.1	1905	12	US-10-282-122A-17125	Sequence 17125, A
	26	25.8	24.1	4544	14	US-10-081-051-58	Sequence 58, Appl
	27	25.8	24.1	7461	9	US-09-764-860-1135	Sequence 1135, App
	28	25.8	24.1	7461	9	US-09-764-860-1135	Sequence 1135, App
	29	25.8	24.1	7461	14	US-10-091-548-91	Sequence 91, Appl
	30	25.8	24.1	7461	14	US-10-074-095-1135	Sequence 1135, Ap
	31	25.8	24.1	7461	15	US-10-212-872-1135	Sequence 1135, Ap
	32	25.8	24.1	32174	9	US-09-908-711-158	Sequence 158, App
	33	25.8	24.1	32174	9	US-09-764-860-1134	Sequence 1134, Ap
	34	25.8	24.1	32174	9	US-09-764-860-1134	Sequence 1134, Ap
	35	25.8	24.1	32174	9	US-09-860-670-232	Sequence 232, App
	36	25.8	24.1	32174	9	US-09-764-904-90	Sequence 90, Appl
	37	25.8	24.1	32174	10	US-09-764-891-6480	Sequence 6480, Ap
	38	25.8	24.1	32174	10	US-09-764-891-10135	Sequence 10135, A
	39	25.8	24.1	32174	10	US-09-764-891-10179	Sequence 10179, A
	40	25.8	24.1	32174	14	US-10-091-548-90	Sequence 90, Appl
	41	25.8	24.1	32174	14	US-10-074-095-1134	Sequence 1134, Ap
	42	25.8	24.1	32174	15	US-10-212-872-1134	Sequence 1134, Ap
	43	25.8	24.1	32174	15	US-10-074-024-609	Sequence 609, App
	44	25.8	24.1	32174	15	US-10-227-646-232	Sequence 232, App
	45	25.8	24.1	32174	15	US-10-242-515-2645	Sequence 2645, Ap

#### ALIGNMENTS

#### RESULT 1

US-10-012-697-280  
; Sequence 280, Application US/10012697  
; Publication No. US20030215803A1  
; GENERAL INFORMATION:  
; APPLICANT: Escobedo, Jaime  
; APPLICANT: Garcia, Pablo Dominguez  
; APPLICANT: Kassam, Altaf  
; APPLICANT: Lamson, George  
; APPLICANT: Scott, Beth  
; APPLICANT: Drmanac, Radoje  
; APPLICANT: Crkvenjakov, Radomir  
; APPLICANT: Dickson, Mark  
; APPLICANT: Drmanac, Snezana  
; APPLICANT: Labat, Ivan  
; APPLICANT: Leshkowitz, Dena  
; APPLICANT: Kita, David  
; APPLICANT: Garcia, Veronica  
; APPLICANT: Jones, Lee William  
; APPLICANT: Stache-Crain, Birgit  
; TITLE OF INVENTION: HUMAN GENES AND GENE EXPRESSION PRODUCTS  
; TITLE OF INVENTION: ISOLATED FROM HUMAN PROSTATE  
; FILE REFERENCE: 2300-16252  
; CURRENT APPLICATION NUMBER: US/10/012,697  
; PRIOR FILING DATE: 2003-01-21  
; PRIOR APPLICATION NUMBER: 60/254,648  
; PRIOR FILING DATE: 2000-12-07  
; PRIOR APPLICATION NUMBER: 60/275,668  
; PRIOR FILING DATE: 2001-03-13  
; NUMBER OF SEQ ID NOS: 1568  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 280  
; LENGTH: 524  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 479, 484, 486, 487, 496, 501, 505, 506, 509, 511, 513, 514,

```

/ LOCATION: 515, 516, 521, 523
/ OTHER INFORMATION: n = A,T,C or G
/
/ NAME/KEY: misc feature
/
/ LOCATION: 479_484, 486, 487, 496, 501, 505, 506, 509, 511, 513, 514,
/ LOCATION: 515, 516, 521, 523
/ OTHER INFORMATION: n = A,T,C or G
/
/ NAME/KEY: misc feature
/
/ LOCATION: 479_484, 486, 487, 496, 501, 505, 506, 509, 511, 513, 514,
/ LOCATION: 515, 516, 521, 523
/ OTHER INFORMATION: n = A,T,C or G
/
/ NAME/KEY: misc feature
/
/ LOCATION: 479_484, 486, 487, 496, 501, 505, 506, 509, 511, 513, 514,
/ LOCATION: 515, 516, 521, 523
/ OTHER INFORMATION: n = A,T,C or G
/
/ NAME/KEY: misc feature
/
/ LOCATION: 479_484, 486, 487, 496, 501, 505, 506, 509, 511, 513, 514,
/ LOCATION: 515, 516, 521, 523
/ OTHER INFORMATION: n = A,T,C or G
/
/ NAME/KEY: misc feature
/
/ LOCATION: 479_484, 486, 487, 496, 501, 505, 506, 509, 511, 513, 514,
/ LOCATION: 515, 516, 521, 523
/ OTHER INFORMATION: n = A,T,C or G
/
/ US-10-012-697-280
/
Query Match 25.6%; Score 27.4; DB 15; Length 524;
Best local Similarity 24.3%; Pred. No. 1.2e+02;
Matches 25; Conservative 9; Mismatches 69; Indels 0; Gaps 0;

```

	Query Match	25.6%	Score 27.4;	DB 15;	Length 524;
	Best Local Similarity	24.3%;	Pred. No. 1.2e+02;		
	Matches 25; Conservative	9;	Mismatches 69;	Indels 0;	Gaps 0;
QY	2	GAAGAAUNNNNUAUGANNNNNNNCUNNNUUUNNNNNNAAAAAANUAAAACUUNNNNCU	61		
DG	232	GAAGAACTCTGTGTAGACATGATGCTTTCTCTATACAGAAGAAAGATACAAACTTGATTGCT	291		
QY	62	UNAGNNNNNNNNNNNNNGAGAAAUNNNNNNNNNNNNNNNNNNNUUAG	104		
DG	292	TCAGTAAAAACATTTTAGAGATTGGTTCCTTTTTGTGAG	334		

```

RESULT 2
US-10-012-697-493
; Sequence 493, Application US/10012697
; Publication No. US20030215803A1
; GENERAL INFORMATION:
; APPLICANT: Escobedo, Jaime
; APPLICANT: Garcia, Pablo Dominguez
; APPLICANT: Kassam, Altaf
; APPLICANT: Lamson, George
; APPLICANT: Scott, Beth
; APPLICANT: Drmanac, Radoje
; APPLICANT: Crkvenjakov, Radomir
; APPLICANT: Dickson, Mark
; APPLICANT: Drmanac, Snezana
; APPLICANT: Labat, Ivan
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Kita, David
; APPLICANT: Garcia, Veronica
; APPLICANT: Jones, Lee William
; APPLICANT: Stache-Crain, Birgit
; TITLE OF INVENTION: HUMAN GENES AND GENE EXPRESSION PRODUCTS
; TITLE OF INVENTION: ISOLATED FROM HUMAN PROSTATE
; FILE REFERENCE: 2300-16252
; CURRENT APPLICATION NUMBER: US/10/012,697
; CURRENT FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: 60/254,648
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: 60/275,668
; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 1568
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 493
; LENGTH: 761
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```

```

; NAME/KEY: misc_feature
; LOCATION: 534, 596, 615, 633, 641, 648, 656, 710, 729, 736, 744, 748,
; LOCATION: 749, 751, 752
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 534, 596, 615, 633, 641, 648, 656, 710, 729, 736, 744, 748,
; LOCATION: 749, 751, 752
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 534, 596, 615, 633, 641, 648, 656, 710, 729, 736, 744, 748,
; LOCATION: 749, 751, 752
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 534, 596, 615, 633, 641, 648, 656, 710, 729, 736, 744, 748,
; LOCATION: 749, 751, 752
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 534, 596, 615, 633, 641, 648, 656, 710, 729, 736, 744, 748,
; LOCATION: 749, 751, 752
; OTHER INFORMATION: n = A,T,C or G
US-10-012-697-493

Query Match          25.6%; Score 27.4; DB 15; Length 761;
Best Local Similarity 24.3%; Pred. No. 1.4e+02;
Matches 25; Conservative 9; Mismatches 69; Indels 0; Gaps

Qy      2   GAAAGUNNNNUNAUGNNNNNNNNNNUUUNNNNNNAAAAAAUAUAAAACAUNNNNCU 61
Db      237 GAAGAATCTTGTGATGACATGAAGTTTCTACTACAGAGAAAGATACAACTTGATTGCT 296

```

Query Match	25.6%	Score 27.4;	DB 15;	Length 761;
Best Local Similarity	24.3%	Pred. No. 1.4e+02;		
Matches 25;	Conservative 9;	Mismatches 69;	Indels 0;	Gaps 0

  

Qy	2	GAAAUUNNNUNUAUGNNNNNNCUNNNUUUNNNNNNAAAAUAUAAAACAUNNNNCU	61
		:         :         :         :         :         :	
Db	237	GAAAACTCTGTTAGACATGATGTTCTCTACAGAGAAAGATACAACTTGATTGCT	296
		:         :         :         :         :         :	
Qy	62	UNAGUNNNNNNNNNNAGAAUUNNNNNNNNNNNNNNNNNNUUAG	104
		:         :         :         :         :         :	
Db	297	TCAGTAAAAACACTTTTAGAGATTGTGTTCTCTTTTGTGTGAG	339
		:         :         :         :         :         :	

RESULT 3  
 US-10-012-697-418  
 / GENERAL INFORMATION:  
 / APPLICANT: Escobedo, Jaime  
 / APPLICANT: Garcia, Pablo Dominguez  
 / APPLICANT: Kasam, Altaf  
 / APPLICANT: Lamson, George  
 / APPLICANT: Scott, Beth  
 / APPLICANT: Drmanac, Radoje  
 / APPLICANT: Crkvenjakov, Radomir  
 / APPLICANT: Dickson, Mark  
 / APPLICANT: Drmanac, Snerzana  
 / APPLICANT: Labat, Ivan  
 / APPLICANT: Leshkowitz, Dena  
 / APPLICANT: Kita, David  
 / APPLICANT: Garcia, Veronica  
 / APPLICANT: Jones, Lee William  
 / APPLICANT: Stache-Crain, Birgit  
 / TITLE OF INVENTION: HUMAN GENES AND GENE EXPRESSION PRODUCTS  
 / TITLE OF INVENTION: ISOLATED FROM HUMAN PROSTATE  
 / FILE REFERENCE: 2300-16252  
 / CURRENT APPLICATION NUMBER: US/10/012,697  
 / CURRENT FILING DATE: 2003-01-21  
 / PRIOR APPLICATION NUMBER: 60/254,648  
 / PRIOR FILING DATE: 2000-12-07  
 / PRIOR APPLICATION NUMBER: 60/275,668  
 / PRIOR FILING DATE: 2001-03-13  
 / NUMBER OF SEQ ID NOS: 1568  
 / SOFTWARE: FastSeq for Windows Version 4.0  
 / OTHER INFORMATION: n = A,T,C or G  
 US-10-012-697-418

Query Match 25.6%; Score 27.4; DB 15; Length 76  
 Best Local Similarity 24.3%; Pred. No. 1.4e-02



US-10-012-697-1307

	Query Match	25.6%	Score 27.4	DB 15	Length 929
	Best Local Similarity	24.3%	pred. No. 1.5e+02		
	Matches	25	Conservative	9	Mismatches 69
					Indels 0
					Gaps 0
QY	2	GAAGAAUNNNNUAUAUNNNNNNNNNUNNNUNNNNNNNAAAAAUAUAACAUNNNNCU	61		
		:           :           :           :			
Db	237	GAAGAACTCTTTGTAGACATGATGTTCTCTATACAGAGAAAGATACAAACTTGATTGCT	296		
QY	62	UNAGNNNNNNNNNNNNAGAAUNNNNNNNNNNNNNNNNNNNNNNNUUAG	104		
		:           :           :           :			
Db	297	TCAGTAAAAACACTTTAGTAGATTGTTGTCCTTTTGTGAG	339		

## RESULT 6

```

US-10-311-455-239/c
; Sequence 239, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detecting
; TITLE OF INVENTION: Cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; CURRENT APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 239
; LENGTH: 6419
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-239

```

Query Match		25.0%;	Score 26.8;	DB 14;	Length 6419;
Best Local Similarity		24.0%;	Pred. No. 4.4e+02;		
Matches	25;	Conservative	9;	Mismatches 70;	Indels 0; Gaps 0;
QY	3	AAAAUNNNNUNUAUGNNNNNNVCUNNNNUUNNNNNNNAAAANAACAAUUNNNNNCUU	62		
dDb	1272	AAAAAATAAAATTATTAATTTCCCATCTTCAACAATAAAAAACTTAAAAATCGAAACTT	1213		
QY	63	NAGNNNNNNNNNNNAGAAAUNNNNNNNNNNNNNNNNNNNUUUAGCA	106		
Dd	1212	AAAAAACAACTTACCACCAAAATCACCCCAACAACTATTTCCTAACCA	1169		

## RESULT 7

```

US-10-311-455-1166/c
; Sequence 1166, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30

```

```

; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1166
; LENGTH: 13814
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated gel
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 9297
; OTHER INFORMATION: n is a o r g o r t
US-10-311-455-1166

```

Query Match	25.0%	Score 26.8;	DB 14;	Length 13814;
Best Local Similarity	22.1%;	Pred. No. 5.9e+02;		
Matches	23;	Conservative 11;	Mismatches 70;	Indels 0; Gaps 0;
Qy	3	AAAUNNNNNUAUAUGNNNNNCUNNNUUUNNNNNNAAAAAUANAACAUNNNNNCUU	62	
Dd	2230	AAATTAATTTTTAAATTTCTTAAACACTTTTAAACAAAAATAATTCCTTAAACACTA	2171	
Qy	63	NAGNNNNNNNNNNNAGAAUUNNNNNNNNNNNNNNNNNNNNNUUJAGA	106	
Dd	2170	AAAAACAAAAATTAATTAATTAATTAATTCCAATCTCTTCCCAATTATCA	2127	

## RESULT 8

```

US-10-312-841-2/C
; Sequence 2, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MH
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 2
; LENGTH: 3673778
;

```

```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (379615)
US-10-312-841-2

```

Query Match 25.0%; Score 26.8; DB 14; Length 3673778;  
Best Local Similarity 22.1%; Pred. No. 9.4e+02;  
Matches 23; Conservative 11; Mismatches 70; Indels 0; G

QY	3	AAAAUNNNNNUNNAUGNNNNNNNNCUNNNUUUNNNNNAAAAAUNAAAAUNNNNNCUU	62
Db	2995750	AAAAATATCTTCATATAAAAAACATAAATTTATCTATAATAATCTCAACACATATCATTA	
QY	63	NAGNNNNNNNNNNAGAAAAUNNNNNNNNNNNNNNNNNNNUUAGCA	106
Db	2995690	AAATATAACATACCCAAAAAATTTAAATTAATATATATTTTAAACA	2995647

## RESULT 9

US-10-027-632, 323248/c  
; Sequence 323248, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

M nucleic - nucleic search, using sw model

Run on: March 8, 2004, 04:24:32 ; Search time 2381 Seconds  
(without alignments)  
1341.980 Million cell updates/sec

Title: US-09-646-651C-2

Perfect score: 107

Sequence: 1 ggaagaaunnnnnnauaun.....nnnnnnnnnnnnuagcag 107

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

result	No.	Score	Match	Length	ID	Description
c	1	29.6	27.7	679	28	BZ043484
c	2	29	27.1	453	10	BG059751
c	3	28.8	26.9	248	13	BX684795
c	4	28.8	26.9	685	28	BH488315

C	5	28.8	26.9	794	28	BH489963
C	6	28.6	26.7	1125	29	CNS057N6
C	7	28	26.2	367	28	AQ135026
C	8	28	26.2	394	28	AQ205663
C	9	28	26.2	394	28	AQ205663
C	10	28	26.2	406	28	AQ211558
C	11	28	26.2	406	28	AQ211558
C	12	27.8	26.0	539	9	AL915410
C	13	27.6	25.8	539	9	AL915411
C	14	27.6	25.8	618	28	BH719067
C	15	27.4	25.6	807	28	BH654909
C	16	27.4	25.6	807	28	BH654909
C	17	27.4	25.6	508	14	CB050688
C	18	27.4	25.6	544	14	CB050532
C	19	27.4	25.6	544	14	CB050531
C	20	27.4	25.6	704	9	AL720964
C	21	27	25.2	234	28	CC397779
C	22	27	25.2	500	9	AU087590
C	23	27	25.2	714	28	BH970555
C	24	27	25.2	730	28	BZ032632
C	25	27	25.2	779	28	CC098521
C	26	27	25.2	810	28	BH480090
C	27	27	25.2	860	12	BG976823
C	28	27	25.2	942	14	CB182382
C	29	27	25.2	1101	29	CNS01776
C	30	26.8	25.0	333	14	CF544518
C	31	26.8	25.0	463	28	AQ466253
C	32	26.8	25.0	540	28	AZ067096
C	33	26.8	25.0	653	29	CE241947
C	34	26.8	25.0	686	29	CE846494
C	35	26.6	24.9	963	29	CNS074X6
C	36	26.6	24.9	429	28	AQ513768
C	37	26.6	24.9	458	28	AQ517382
C	38	26.6	24.9	487	28	BZ173324
C	39	26.6	24.9	719	29	AG008606
C	40	26.6	24.9	739	14	CF259384
C	41	26.6	24.9	781	28	CC077740
C	42	26.6	24.9	841	28	BZ262911
C	43	26.6	24.9	857	29	CNS01053
C	44	26.6	24.9	859	28	CC198967
C	45	26.4	24.7	182	28	BH911389
C	46	26.4	24.7	418	12	BG941688

## ALIGNMENTS

RESULT 1  
BZ043484/c  
LOCUS  
DEFINITION  
BZ043484  
ACCESION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

BZ043484  
lkgs57f11.bl B.oleracea002 Brassica oleracea Genomic, genomic survey  
sequence.

BZ043484  
BZ043484.1 GI:23635947  
GSS.  
Brassica oleracea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
1 (bases 1 to 679)

Delehaunty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T.,  
Nash,W., Rabinowicz,P.D. and Wilson,R.K.  
Whole genome shotgun reads from Brassica oleracea  
Unpublished (2002)  
Contact: Richard K. Wilson  
Genome Sequencing Center  
Washington University School of Medicine  
Email: submissions@watson.wustl.edu

Seq primer: -21UPOT forward  
Class: shotgun  
High quality sequence start: 17  
High quality sequence stop: 551.

```

Best Local Similarity   23.3%;   Pred. No. 1.7e+03;
Matches    24; Conservative    11; Mismatches    68; Indels      0; Gaps      0;

QY      1  GGAAAUUNNNNNUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAACAUNNNNC 60
||||| : : : : : | : : : : : ||||| : : : : :
Db      189 GGAGAAATGAAATTAACTTTTCCTCTCGATTAGGAGAAAAACAATGGAAACATTTTAAA 248
||||| : : : : : | : : : : : ||||| : : : : :

QY      61 UUNAGNNNNNNNNNNNAGAAUUNNNNNNNNNNNNNNUUA 103
||||| : : : : : | : : : : : ||||| : : : : :
Db      249 ACAAGTTTATATCTGAGAGAAAAAAGAAGTAATCATTCATA 291
||||| : : : : : | : : : : : ||||| : : : : :

RESULT 3
EX684795/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Sanger Xenopus tropicalis EST project 2001 (11_2003)
Contact: Croning MDR
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk

This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
cDNA was oligo dt primed from Sug of poly A+ RNA from neurula.
EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the
5' end and NotI at the 3' end.
Vector: pCS107; Site_1: EcoRI; Site_2: NotI
Host: Escherichia coli DH10B.
Location/Qualifiers
1..248
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clones="TNeu075o20"
/dev_stage="neurula"
/lab_host="Escherichia coli DH10B"
/clone_lib="XGC-neurula"
/note="vector: pCS107; Site_1: EcoRI; Site_2: NotI; cdna
was oligo dt primed from Sug of poly A+ RNA from neurula.
EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end."

ORIGIN

Query Match          26.9%; Score 28.8; DB 13; Length 248;
Best Local Similarity 31.7%; Pred. No. 2.3e+03;
Matches    26; Conservative     8; Mismatches    48; Indels      0; Gaps      0;

QY      3  AAAAUUNNNNNUAUGNNNNNNCUNNNUUUNNNNNNAAAAANUANAACAUNNNNCUU 62
||||| : : : : : | : : : : : ||||| : : : : :
Db      110 AAATGATGTGAAGAGTTTACCCTTTTTTCAAAAAAAAAAAAAAATTTCAACGG 51
||||| : : : : : | : : : : : ||||| : : : : :

QY      63 NAGNNNNNNNNNNNNAGAAU 84
||||| : : : : : | : : : : : ||||| : : : : :
Db      50 GAGTAAAAAATAAAAAAAT 29
||||| : : : : : | : : : : : ||||| : : : : :

RESULT 4
BH488315
```

```

DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: IR
Class: sheared ends.
Location/Qualifiers
    1..794
       /organism="Brassica oleracea"
       /mol_type="genomic DNA"
       /strain="TO1000DH3"
       /db_xref="taxon:3712"
       /clone="BOKA57"
       /clone_lib="BOKX"
       /note="Vector: pHOs1; Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into pHOs1 using BstXI linkers"

FEATURES
     source
         Query Match          26.9%; Score 28.8; DB 28; Length 794;
         Best Local Similarity 24.5%; Pred. No. 1.5e+03;
         Matches 26; Conservative 10; Mismatches 70; Indels 0; Gaps 0;

ORIGIN
      QY      1  GGAAAAUUNNNNNUAUAGUNNNNNNNCUNNNUUUNNNNNNAAAAANTANAACAAUNNNNC   60
                :|:::
      Db     128  GGAAATAATATATTAPAGAGAATTTTITTTTATAGCAATGCAAAATTCACACAATTTTC   69

```

```

QY      61  UUNAGNNNNNNNNNNNAGAAUUNNNNNNNNNNNNNNNNNNNNNNUAGCA 106
      ::      |||||:
Ddb      68  TTCCATCAACAATTACTCAGAAATAACCAAGCTGCTCTCGCATCA 23

RESULT 6
CNS057N6//c
LOCUS
DEFINITION
Tetraodon nigroviridis genome survey sequence T7 end of Clone
010K04 of library A from Tetraodon nigroviridis, genomic survey
sequence
1125 bp      DNA      linear      GSS 26-JUL-2000
CNS057N6

```

AL324843.1 GI:9557725  
GSS: genome survey sequence.  
Tetraodon nigroviridis  
Tetraodon nigroviridis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
Tetraodontoidea; Tetraodontidae; Percomoda.  
1  
REFERENCE  
Roest Crolius H., Jaillon O., Dasilva C., Bouneau L., Fisher C.,  
Bernot A., Fizanes C., Wincker P., Brottier P., Quetier F.,  
G...

genome-wide analysis  
 using Tetraodon nigroviridis DNA sequence  
 Nat. Genet. 25 (2), 235-238 (2000)  
 20294633  
 10835645  
 2  
 REFERENCES  
 2  
 AUTHORS  
 Roest Crolius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,  
 Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,  
 Saurin,W., Bernot,A. and Weissenbach,J.  
 Characterization and repeat analysis of the compact genome of the  
 freshwater pufferfish Tetraodon nigroviridis  
 Genome Res. 10 (7), 939-949 (2000)  
 20359837  
 10893143  
 3 (bases 1 to 1125)  
 Genoscope.  
 Direct Submission  
 Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :  
 BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr  
 - Web : www.genoscope.cns.fr)

- web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
 This sequence is a single read and was generated as part of a large  
 scale clone-end sequencing project of the Tetraodon nigroviridis  
 genome. For more information, please take a look at  
<http://www.genoscope.cns.fr/Tetraodon>.  
 Location/Qualifiers

```

/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:9983"
/cloned_lib="A"
/note="Genoscope sequence ID : C0AA01BF02C3-end : T7"

ORIGIN
Query Match      26.7%; Score 28.6; DB 29; Length 1125;
Best Local Similarity 21.8%; Pred.No. 1.5e+03;
Matches 22; Conservative 12; Mismatches 67; Indels 0; Gaps 0;

QY 3 AAAAAUNNNNUNAUGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNUU 62
    |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:
Db 247 AAAATACTAAATATTTTACCTGCACTTATATATCAAGACCAATTAACCTTCIT 188
    |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:
QY 63 NAGNNNNNNNNNNNNNAGAAUNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNUU 103
    |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:
Db 187 AGCTTTTATCATGATGAACAAGAAGTGATTGGAGTCTGTGTTA 147
    |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:

RESULT 7
AQ135026          367 bp   DNA       linear     GSS 24-SEP-1998
LOCUS             HS 3053.A2 A11.MR CIT Approved Human Genomic Sperm Library D Homo
DEFINITION        sapiens genomic clone Plate=3053 Col=22 Row=A, genomic survey
sequence.
ACCESSION         AQ135026 GI:3526392
VERSION           AQ135026
KEYWORDS          GSS.
SOURCE            Homo sapiens (human)
ORGANISM          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE         1 (bases 1 to 367)
AUTHORS           Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
TITLE             Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
JOURNAL           Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE          99380589
PUBMED           1049764
COMMENT           Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 3053 row: A column: 22
Class: BAC ends
High quality sequence stop: 367.
Location/Qualifiers
1..367
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/cloned_lib="Plate=3053 Col=22 Row=A"
/sex="male"
/cloned_lib="CIT Approved Human Genomic Sperm Library D"
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in
E-Coli DH10B"

FEATURES
source
Query Match      26.2%; Score 28; DB 28; Length 367;
Best Local Similarity 23.5%; Pred.No. 3.1e+03;
Matches 24; Conservative 10; Mismatches 68; Indels 0; Gaps 0;

ORIGIN
Query Match      26.2%; Score 28; DB 28; Length 367;
Best Local Similarity 23.5%; Pred.No. 3.1e+03;
Matches 24; Conservative 10; Mismatches 68; Indels 0; Gaps 0;

QY 2 GAAAAUNNNNUNAUGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCU 61
    |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:
Db 233 GAAAATCTCTTTAAATATATCTGCTTAAAGCAAAGATTAGTACATGCATCC 292
    |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:
QY 62 UNAGNNNNNNNNNNNNNAGAAUNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNUU 103
    |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:
Db 293 TAAGTTTTATAATGTGTAGTAGTAAGTAAAGTGTATTATAACCATA 334
    |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:

RESULT 9
AQ268612          394 bp   DNA       linear     GSS 27-APR-1999
LOCUS             RPCI11-7LJ23.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-7LJ23,
DEFINITION        genomic survey sequence.
ACCESSION         AQ268612 GI:3796216
VERSION           AQ268612
KEYWORDS          GSS.
SOURCE            Homo sapiens (human)

```





```

/organism="Brassica oleracea"
/mol_type="genomic DNA"
/_strain="Tc1000DH3"
/db_xref="taxon:3712"
/_clone="BOHTZ17"
/_clone_lib="BO_2_3 KB"
/_note="Vector: pHO51; Site 1:
genomic DNA inserted into pHO51"

```

RIGIN

[illegible]

RESULT 15	CB050688	204 bp	mRNA	linear	EST 17-JAN-2003
OCUS	CB050688				
DEFINITION	NISC_gj13b02.y1 NCI_CGAP_Pz28 Homo sapiens cDNA clone IMAGE:3272018 5', mRNA sequence.				
ACCESSION	CB050688				
VERSION	CB050688.1	GI:27788975			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 204)				
TITLE	NCI-CGAP <a href="http://www.ncbi.nlm.nih.gov/ncicgap">http://www.ncbi.nlm.nih.gov/ncicgap</a> . National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index				
JOURNAL	Unpublished (1997)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:cgapbs-remail.nih.gov">cgapbs-remail.nih.gov</a>				

cDNA Library Preparation:  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL  
DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC)  
Clone distribution: NCI-CCAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
info@image.llnl.gov  
Plate: LLAM8010 row: D column: 3  
Seq primer: M13RP1 reverse primer (ABI).

seq primer: M13KPI reverse primer (ABI).  
Location/Qualifiers

FEATURES	SOURCE
1. <i>Age</i>	1. <i>Age</i>
2. <i>Gender</i>	2. <i>Gender</i>
3. <i>Marital Status</i>	3. <i>Marital Status</i>
4. <i>Education</i>	4. <i>Education</i>
5. <i>Income</i>	5. <i>Income</i>
6. <i>Occupation</i>	6. <i>Occupation</i>
7. <i>Religion</i>	7. <i>Religion</i>
8. <i>Political Affiliation</i>	8. <i>Political Affiliation</i>
9. <i>Health Status</i>	9. <i>Health Status</i>
10. <i>Travel History</i>	10. <i>Travel History</i>
11. <i>Family Size</i>	11. <i>Family Size</i>
12. <i>Home Ownership</i>	12. <i>Home Ownership</i>
13. <i>Employment Status</i>	13. <i>Employment Status</i>
14. <i>Language Spoken</i>	14. <i>Language Spoken</i>
15. <i>Religious Beliefs</i>	15. <i>Religious Beliefs</i>
16. <i>Political Views</i>	16. <i>Political Views</i>
17. <i>Health Insurance</i>	17. <i>Health Insurance</i>
18. <i>Travel Insurance</i>	18. <i>Travel Insurance</i>
19. <i>Family Income</i>	19. <i>Family Income</i>
20. <i>Home Address</i>	20. <i>Home Address</i>
21. <i>Employment History</i>	21. <i>Employment History</i>
22. <i>Language Proficiency</i>	22. <i>Language Proficiency</i>
23. <i>Religious Practices</i>	23. <i>Religious Practices</i>
24. <i>Political Participation</i>	24. <i>Political Participation</i>
25. <i>Health Care Usage</i>	25. <i>Health Care Usage</i>
26. <i>Travel Preferences</i>	26. <i>Travel Preferences</i>
27. <i>Family Structure</i>	27. <i>Family Structure</i>
28. <i>Home Environment</i>	28. <i>Home Environment</i>
29. <i>Employment Satisfaction</i>	29. <i>Employment Satisfaction</i>
30. <i>Language Usage</i>	30. <i>Language Usage</i>
31. <i>Religious Community</i>	31. <i>Religious Community</i>
32. <i>Political Engagement</i>	32. <i>Political Engagement</i>
33. <i>Health Care Access</i>	33. <i>Health Care Access</i>
34. <i>Travel Itinerary</i>	34. <i>Travel Itinerary</i>
35. <i>Family Dynamics</i>	35. <i>Family Dynamics</i>
36. <i>Home Security</i>	36. <i>Home Security</i>
37. <i>Employment Stability</i>	37. <i>Employment Stability</i>
38. <i>Language Fluency</i>	38. <i>Language Fluency</i>
39. <i>Religious Observance</i>	39. <i>Religious Observance</i>
40. <i>Political Involvement</i>	40. <i>Political Involvement</i>
41. <i>Health Care Costs</i>	41. <i>Health Care Costs</i>
42. <i>Travel Expenses</i>	42. <i>Travel Expenses</i>
43. <i>Family Well-being</i>	43. <i>Family Well-being</i>
44. <i>Home Amenities</i>	44. <i>Home Amenities</i>
45. <i>Employment Benefits</i>	45. <i>Employment Benefits</i>
46. <i>Language Skills</i>	46. <i>Language Skills</i>
47. <i>Religious Values</i>	47. <i>Religious Values</i>
48. <i>Political Ideology</i>	48. <i>Political Ideology</i>
49. <i>Health Care Quality</i>	49. <i>Health Care Quality</i>
50. <i>Travel Destinations</i>	50. <i>Travel Destinations</i>
51. <i>Family Relationships</i>	51. <i>Family Relationships</i>
52. <i>Home Location</i>	52. <i>Home Location</i>
53. <i>Employment Opportunities</i>	53. <i>Employment Opportunities</i>
54. <i>Language Proficiency</i>	54. <i>Language Proficiency</i>
55. <i>Religious Beliefs</i>	55. <i>Religious Beliefs</i>
56. <i>Political Views</i>	56. <i>Political Views</i>
57. <i>Health Insurance</i>	57. <i>Health Insurance</i>
58. <i>Travel Insurance</i>	58. <i>Travel Insurance</i>
59. <i>Family Income</i>	59. <i>Family Income</i>
60. <i>Home Address</i>	60. <i>Home Address</i>
61. <i>Employment History</i>	61. <i>Employment History</i>
62. <i>Language Proficiency</i>	62. <i>Language Proficiency</i>
63. <i>Religious Practices</i>	63. <i>Religious Practices</i>
64. <i>Political Participation</i>	64. <i>Political Participation</i>
65. <i>Health Care Usage</i>	65. <i>Health Care Usage</i>
66. <i>Travel Preferences</i>	66. <i>Travel Preferences</i>
67. <i>Family Structure</i>	67. <i>Family Structure</i>
68. <i>Home Environment</i>	68. <i>Home Environment</i>
69. <i>Employment Satisfaction</i>	69. <i>Employment Satisfaction</i>
70. <i>Language Usage</i>	70. <i>Language Usage</i>
71. <i>Religious Community</i>	71. <i>Religious Community</i>
72. <i>Political Engagement</i>	72. <i>Political Engagement</i>
73. <i>Health Care Access</i>	73. <i>Health Care Access</i>
74. <i>Travel Itinerary</i>	74. <i>Travel Itinerary</i>
75. <i>Family Dynamics</i>	75. <i>Family Dynamics</i>
76. <i>Home Security</i>	76. <i>Home Security</i>
77. <i>Employment Stability</i>	77. <i>Employment Stability</i>
78. <i>Language Fluency</i>	78. <i>Language Fluency</i>
79. <i>Religious Observance</i>	79. <i>Religious Observance</i>
80. <i>Political Ideology</i>	80. <i>Political Ideology</i>
81. <i>Health Care Costs</i>	81. <i>Health Care Costs</i>
82. <i>Travel Expenses</i>	82. <i>Travel Expenses</i>
83. <i>Family Well-being</i>	83. <i>Family Well-being</i>
84. <i>Home Amenities</i>	84. <i>Home Amenities</i>
85. <i>Employment Benefits</i>	85. <i>Employment Benefits</i>
86. <i>Language Skills</i>	86. <i>Language Skills</i>
87. <i>Religious Values</i>	87. <i>Religious Values</i>
88. <i>Political Involvement</i>	88. <i>Political Involvement</i>
89. <i>Health Care Quality</i>	89. <i>Health Care Quality</i>
90. <i>Travel Destinations</i>	90. <i>Travel Destinations</i>
91. <i>Family Relationships</i>	91. <i>Family Relationships</i>
92. <i>Home Location</i>	92. <i>Home Location</i>
93. <i>Employment Opportunities</i>	93. <i>Employment Opportunities</i>
94. <i>Language Proficiency</i>	94. <i>Language Proficiency</i>
95. <i>Religious Beliefs</i>	95. <i>Religious Beliefs</i>
96. <i>Political Views</i>	96. <i>Political Views</i>
97. <i>Health Insurance</i>	97. <i>Health Insurance</i>
98. <i>Travel Insurance</i>	98. <i>Travel Insurance</i>
99. <i>Family Income</i>	99. <i>Family Income</i>
100. <i>Home Address</i>	100. <i>Home Address</i>

```

1. .204
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3272018"
/sex="male"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="NCI CGAP Pr28"
/notes="Organ: prostate; Vector: p7T73D-Pac (Pharmacia)
with a modified polylinker; Plasmid DNA from the
normalized library NCI CGAP Pr22 was prepared, and ss
circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (clonedIDs
985608-986759, 1101192-1101959, and 1217928-1220615).
Subtraction by Bento Soares and M. Fatima Bonaldo."

```

RIGIN

Query Match 25.6%; Score 27.4; DB 14; Length 204;

```

Best Local Similarity 24.3%; Pred. No. 5.3e+03;
Matches 25; Conservative 9; Mismatches 69; Indels 0; Gaps 0;

QY 2 GAAAAUNNNNNNAUGNNNNNNNNUNNNNNNNAAAAAUAUAAAAAUNNNNCU 61
    ||||| : : : : : : : : : : : : : : : : : : : : : :
Db 99 GAAAACTCTGTTAGACATGATGTTTCCTATACAGAAGAAAGATACAAACTGATTGCT 158

QY 62 UNAGNNNNNNNNNNNNNAGAAUNNNNNNNNNNNNNNNNNNUAG 104
    : : : : : : : : : : : : : : : : : : : : : :
Db 159 TCAGTAAACACACTTTTACAGATGTGTGTTCTCTTTTGTGAG 201
    : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: March 8, 2004, 06:00:27  
Job time : 2387 secs